SOME APPLICATIONS OF HOLOGRAPHY

TO IMAGE PROCESSING

A thesis presented for the degree of

Doctor of Philosophy

in Electrical Engineering

in the University of Canterbury

Christchurch, New Zealand.

by

P.T. Gough B.E.(Hons)

1974

ABSTRACT

A successful implementation is reported of a simple and rapid method of constructing computer generated holograms, by overprinting the standard line printer characters. The theoretical and practical limitations of computer generated holograms are discussed with reference to a number of optical and computational experiments reported here.

A new technique, called "speckle holography", is described and its successful simulation in the optical laboratory is reported. This technique is an extension of Labeyrie's speckle interferometry and it permits true images to be obtained through randomly fluctuating media. Comments are also made on Knox and Thompson's extension of speckle interferometry for which it is claimed that it is not necessary to use a holographic approach to obtain a true image. The relevance of this latter extension to future astronomical practice is assessed.

ACKNOWLEDGEMENTS

I am grateful to my supervisor Dr R.H.T. Bates, for his encouragement and enthusiastic assistance during the course of this project. Most of the time we were in agreement.

I must also thank Dr T.M. Peters and Dr P.R. Smith who provided some of the biomedical data and some of their computer programs. I am grateful to Dr K.A. Haines and C. Miles for assistance in the optical part of this project, and to W.K. Kennedy and Dr T.M. Peters for the use of their image scanner and photographic facsimile machine.

I also wish to thank the University Grants Committee for some financial assistance and finally my wife Janet who provided money and encouragement, both of which at times I needed.

TABLE OF CONTENTS

GLOSSARY

In this thesis symbols used have the following meanings unless otherwise indicated in the text. Wherever possible object and image distributions are denoted by small letters and their equivalent FT distributions are denoted by large letters.

$a(x,y)$	the psf of a lens with a limited diameter and no aberrations
$A(\alpha,\beta)$	the FT of $a(x,y)$
$b(x,y)$	the amplitude distribution of the best available image of a diffraction limited coherent optical system
$c(x,y)$	the intensity distribution of the best available in a diffraction limited incoherent optical system
cgh	computer generated hologram/holography
d	density of a developed film (Appendices A,B)
$d(x,y)$	the intensity of the psf
$E(\alpha,\beta)$, $\underline{E}(\alpha,\beta)$	the speckle transforms of two unresolvable stars
f	focal length of a lens
$g(x,y)$	the amplitude distribution of an image
$h(x,y)$	the psf of an optical system
$i(x,y)$	the intensity distribution of an image
$i_m(x,y)$	the speckle interferogram
j	preceding an imaginary number
k, ℓ, m, n	are integers

$k(x_o,x_i,y_o,y_i,\lambda)$	the transmission function of an optical system from the x_o,y_o plane to the x_i,y_i plane
$\bar{k}(x_o,x_i,y_o,y_i)$	the transmission function evaluated at $\bar{\lambda}$
M	magnification of a lens system (Chapter 1)
M	large integer (Chapters 5,6)
N	large integer
$o(x,y)$	object distribution
$O(\alpha,\beta)$	generalised object wavefront in the α,β plane. This is usually the FT distribution of the object $o(x,y)$ although it may be its Fresnel distribution
$p(x,y)$	the perfect image amplitude distribution
psi/v	point spread invariant/variant
psf	point spread function
$Q(\alpha,\beta)$	ideal inverse filter
R	minimum resolvable detail in the object
$R(\alpha,\beta)$	reference wavefront
r,s	cartesian coordinate system
r,θ	polar coordinate system
S	separation between two parts of an object
$	angular diameter of the seeing disc
$t(\alpha,\beta)$	transmissivity of a transparency in the α,β plane
ΔT	exposure time of a film
$\hat{V}(\alpha,\beta)$	speckle transform of a resolvable star
$\hat{v}(x,y)$	autocorrelation of the best available image of a resolvable star. The FT of $\hat{V}(\alpha,\beta)$
W	total width of an extended stellar object

x, y	generalised cartesian coordinates. Hence
x_o, y_o	is the object plane,
x_ℓ, y_ℓ	is the lens plane,
x_i, y_i	is the image plane.
α, β	cartesian coordinates in the diffraction/filter/ hologram plane
ξ, η	cartesian coordinate system the same as the α, β system but rotated by θ
Δ, Δ^+	small displacements
γ	the gamma or contrast of a film
$\gamma(P_1, P_2, \tau)$	the complex degree of coherence between two points P_1, P_2 at times separated by τ
$\Gamma(P_1, P_2, \tau)$	the mutual coherence function between P_1, P_2
λ	the wavelength
$\bar{\lambda}$	the mean wavelength
$\Delta\lambda$	the bandwidth
ρ, ϕ	polar coordinate system
ζ	diameter of a lens
$\otimes$	convolution
$\star$	correlation
$*$	superscript denoting complex conjugate
$'$	superscript denoting differentiation
$J_n(\xi)$	Bessel function
$\lvert\ \rvert$	modulus
FT	Fourier transform
$\longleftrightarrow$	FT pairs
∞	infinity
$\propto$	proportional

PREFACE

Image processing is important in all branches of science and engineering where a picture or an image is used to evaluate the result of an experiment. We try to obtain the best reproduction of the real object that is possible. An image produced by a low quality imaging system is often inconclusive, so we must investigate ways to produce the best possible image in any given situation.

This thesis reports on research into some particular ways to improve the images formed by three different imaging systems. The first example is in biomedical engineering, and is an attempt to image a cross-section of an optically opaque body using data gathered from a series of X-ray photographs. The second example is an attempt to correct a photograph taken from a camera which moved during the exposure. The third example details a new technique where it is possible to image a certain class of objects through a grossly distorting medium in front of the imaging system. This new technique has application in astronomy.

As yet, the processing of images does not fit neatly into any one scientific discipline. Therefore, it is not so strange that an engineer with a background in microwave radiation should deal with these problems, since all these systems use electromagnetic radiation to form the image. In common with most interdisciplinary problems there are few accepted conventions so Chapter 1 outlines the general principles of image formation and introduces most of the conventions used throughout this thesis.

Chapter 1 discusses the coherence of sources, imaging, the Fourier transform, the importance of the concept of spatial frequency in image processing, and also includes a short historical review of some important advances in image processing.

An important tool in image processing is holography which is now used in a variety of imaging applications, including three-dimensional 'photographs', spatial filtering, small displacement measurements and microwave antenna measurements. In all of these fields, situations arise where it is usually easier to measure the radiation pattern from a radiating object, rather than the details of the object itself. The properties of the radiation that we wish to measure are the modulus and the phase of a wavefront as it crosses a particular surface in space, and in most cases the measurement of the phase is the more difficult problem. If we require the phase to be measured directly, the response of the detectors across the wavefront must be shorter than the time period of the wavefront. Although this may be practical for some low frequency microwave wavefronts, it is more usual for the detector to take many time periods to detect the presence of the wave. It is for the latter situation that holography is used to measure both the modulus and the phase of the radiated wavefront. Chapter 2 outlines the principles of holography and includes some examples of the way holograms are formed by radiated electromagnetic and acoustic waves. Included in this chapter is a section on holograms not formed by radiated waves, but formed artificially in a digital computer. The computer produces a "synthetic aperture hologram" that may be used as

though it had been constructed optically. The various types of computer generated holograms and their particular advantages are explained.

The original research is reported in Chapters 3, 4 and 5. Chapter 3 outlines the way in which a clear cross-section of an opaque body may be determined from a series of X-ray photographs all exposed at slightly different angles to the body. This process is implemented using a particular type of computer generated hologram and the construction of this hologram is detailed, since the problems encountered are typical of all computer generated holograms.

Chapter 4 reports on an attempt to use a computer generated hologram as a spatial filter to improve a blurred photograph. This chapter highlights the practical limitations of computed holograms.

Chapter 5 outlines a holographic approach to imaging a radiating object through a distorting medium close to the imaging lens. This holographic technique is applicable to astronomy where a telescope views a group of stars through the earth's atmosphere. It is shown how a series of images taken by the telescope may provide sufficient information to correct the distorting affect of the atmosphere and produce an image that is, in principle, distortion-free. A laboratory simulation of this process is reported and some images, before and after processing, are illustrated. Since the work reported in this thesis was completed, claims have been made that a technique has been developed that suggests that it is not necessary to use holography to produce a distortion free image of a stellar object. Chapter 6 deals with this new process and discusses possible difficulties with its practical application.

Publications on topics relevant to this thesis are:

GOUGH P.T. and BATES R.H.T. (1972) "Computer Generated Holograms for Processing Radiographic Data", Compt. Biomed. Res., Vol. 5, 700-708.

BATES R.H.T., GOUGH P.R. and PETERS T.M. (1974) "Some developments in forming Holograms without coherent light", Digest of Papers, Intnl. Opt. Computing Conf., Zurich, Switzerland (IEEE Catalog No. 74 CHO862-3C), pp100-102.

BATES R.H.T., GOUGH P.T. and NAPIER P.J. (1973) "Speckle Interferometry gives Holograms of Multiple Star Systems", Astron. & Astrophys., Vol. 22, pp319-320.

GOUGH P.T. and BATES R.H.T. (1974) "Speckle Holography", Opt. Act., Vol. 21, pp243-254.

BATES R.H.T. and GOUGH P.T. (1974) "On optical processing of radiation received from incoherent objects viewed through random media", Submitted to IEEE Trans. Compt. (Special Issue on Optical Computing).

GOUGH P.T. (1972) "Computation of Accurate Cross-sections of Limbs from Standard X-ray plates". Proc. 5th Hawaii Int. Conf. on System Sciences, pp 241-243.

CHAPTER 1

IMAGING

1.1 INTRODUCTION

An image may be defined as an artificial imitation of the external form of an object. However, for the purposes of this thesis, image formation is restricted to plane objects. This then includes real images formed by light, photographs, transparencies and visual computer printouts, but excludes statues etc.

The oldest imaging system known is undoubtedly the eye. A collection of light waves enters the eye, and is focused onto the retina. Electrical impulses are sent to the brain from the retina, and from these, we interpret which objects produced the original waves.

An imaging system then, is a process that forms a facsimile of a particular object. It can be called 'perfect', if the facsimile, or image, is indistinguishable from the object. Many imaging systems are used to scale the object. That is, the image is 'perfect', except that it is smaller or larger than the object. In most applications, the relative size of the object and image is of little importance, and it is the fidelity of the imaging system that is the criteria of quality. Because of this, the term a 'perfect image' includes images that are larger or smaller than the object, but which are otherwise indistinguishable.

The amount of scaling or magnification between the object and its image is determined by the geometry of the imaging system.

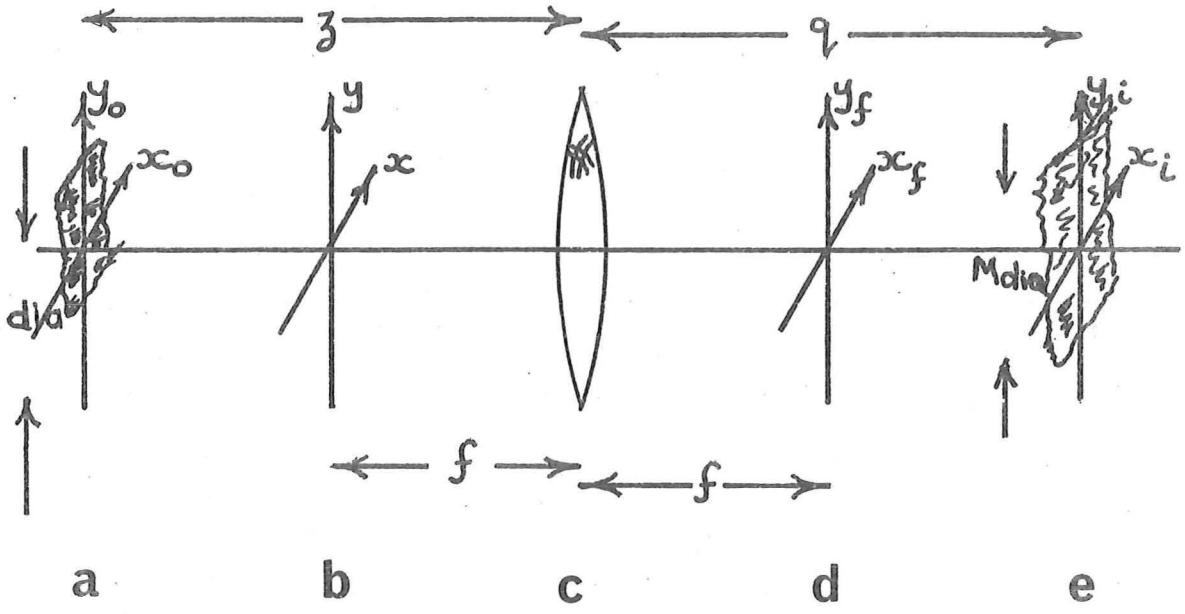


Fig. 1.1 A simple imaging system.
a) Object plane. b) Front focal plane. c) Lens.
d) Rear focal plane. e) Image plane.

The general formula for the position of an image formed by a simple fixed-focus lens of focal length f is (Born & Wolf, 1959)

$$\frac{1}{f} = \frac{1}{z} + \frac{1}{q} \qquad (1.1)$$

where z and q are the distances from the centre of the lens to the object and image planes (Fig. 1.1). The magnification of the image is then

$$M = \frac{q}{z} \qquad (1.2)$$

Thus, if the linear extent of the object is dia, then the linear extent of the image is Mdia.

If the object is represented by $o(x_o, y_o)$ then the perfect image $p(x_i, y_i)$ is defined by

$$p(x_i,y_i) = \frac{1}{M}\, o(Mx_o,My_o) \qquad (1.3)$$

The object is usually a collection of small elementary radiators arranged so that they represent $o(x_o,y_o)$. We call this a distribution of sources. The radiating light from each source is passed through the lens and brought to focus in the image plane. Thus, the image is a collection of small elementary images of the source radiators. We call this a distribution of light. The image distribution of one lens may act as an object distribution to a second lens so that object distributions include distributions of light, as well as physical radiators. If the object source radiators have an amplitude distribution of $o(x,y)$, then the image is also an amplitude distribution given by eqn (1.3). Quite often we only know the object's intensity distribution, that is $|o(x,y)|^2$, in which case the image is also given as an intensity distribution. Thus, we have four types of distributions: object amplitude, image amplitude, object intensity and image intensity. These distributions are not to be confused with the distributions of delta functions as used by many mathematicians.

1.2 THE COHERENCE OF SOURCES

The exact image formed by an optical system depends on the type of radiating sources which constitute the object distribution.

Consider the complex fields at two points in space denoted by $U(P_1,t)$ and $U(P_2,t)$. The cross-correlation of these fields defines the mutual coherence function $\Gamma(P_1,P_2,\tau)$:

$$\Gamma(P_1,P_2,\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T U(P_1,t)\ U^*(P_2,t+\tau)\,dt \qquad (1.4)$$

For convenience, this is normalised to give the complex degree of coherence $\gamma(P_1,P_2,\tau)$:

$$\gamma(P_1,P_2,\tau) = \frac{\Gamma(P_1,P_2,\tau)}{\sqrt{\Gamma(P_1,P_1,0)\Gamma(P_2,P_2,0)}} \qquad (1.5)$$

$$0 \leqslant |\gamma(P_1,P_2,0)| \leqslant 1 \qquad (1.6)$$

The spatial coherence function of the sources in an object is defined as $\gamma(P_1,P_2,0)$ where P_1 and P_2 are points anywhere in the object. If $|\gamma(P_1,P_2,0)|$ is effectively unity for all P_1,P_2 in the object distribution, then the object is termed spatially coherent. If $|\gamma(P_1,P_2,0)|$ is effectively zero everywhere, $P_1 \neq P_2$, then the object is termed spatially incoherent.

The temporal coherence function of any point in an object is $\gamma(P_1,P_1,\tau)$ where τ lies between 0 and ∞. Note that:

$$|\gamma(P_1,P_1,0)| = 1.0,$$

$$|\gamma(P_1,P_1,\infty)| = 0.0.$$

Define the coherence time τ_c, such that,

$$|\gamma(P_1,P_1,\tau_c)| = 0.5 \qquad (1.7)$$

then τ_c represents the time within which a series of wavefronts passing P_1 may be termed coherent.

Consider two points, P_1 and P_2, that lie along the same ray path from the source. Let the distance between them be r. If

$$r \leqslant c\,\tau_c \tag{1.8}$$

where c is the speed of light, then P_1 and P_2 are spatially coherent when τ is zero. Define the coherence distance as r_c, given by

$$r_c = c\,\tau_c. \tag{1.9}$$

1.3 SPATIALLY COHERENT IMAGING

Consider a spatially coherent object. The image produced by an optical system is characterised by the transmission function $k(x_o, x_i, y_o, y_i, \lambda)$ of the system, which is the complex amplitude, per unit area of the object plane, at the point (x_i, y_i), due to radiation of unit amplitude and zero phase at the object point (x_o, y_o) (Born and Wolf, 1959). The transmission function is defined for a single radiated frequency but a spatially coherent object usually has a narrow bandwidth, although this is not absolutely necessary (Papoulis, 1968), If

$$\frac{\Delta\lambda}{\bar{\lambda}} << 1.0 \tag{1.10}$$

where $\Delta\lambda$ is the bandwidth and $\bar{\lambda}$ is the mean wavelength of the radiation, then the radiation may be termed quasimonochromatic.

If the object radiation is quasimonochromatic, the transmission function may be evaluated at the mean wavelength $\bar{\lambda}$. (Often the wavelength dependence is left out of the transmission function as a generally accepted convention.)

Let $o(x_o,y_o)$ represent the amplitude of the radiating sources in the object plane. The image distribution that this object produces in the image plane is $g(x_i,y_i)$ given by:

$$g(x_i,y_i) = \int_0^\infty\iint_{-\infty}^{\infty} o(x_o,y_o)k(x_o,x_i,y_o,y_i,\lambda)d\lambda dx_o dy_o \tag{1.11}$$

If the radiating sources obey the quasimonochromatic condition, eqn (1.11) may be evaluated at the mean wavelength $\bar{\lambda}$.

$$g(x_i,y_i) = \iint_{-\infty}^{\infty} o(x_o,y_o)k(x_o,x_i,y_o,y_i,\bar{\lambda})dx_o dy_o \tag{1.12}$$

The quasimonochromatic transmission function is usually called the point spread function (psf) of the optical system, $\bar{k}(x,y)$, which is defined by

$$\bar{k}(x_o,x_i,y_o,y_i) = k(x_o,x_i,y_o,y_i,\bar{\lambda}) \tag{1.13}$$

The psf may be regarded as the amplitude of the radiation in the image plane due to a point source in the object plane. If the form of the psf is independent of the position of the point source in the object plane then the psf is termed point spread invariant (psi).

In a corrected optical system, the psf is only psi within a certain area around the optical axis. This area is termed the isoplanatic region, and for all object and image

distributions that are contained within this region

$$\bar{k}(x_o,x_i,y_o,y_i) = \bar{k}(x_o-x_i,y_o-y_i) \tag{1.14}$$

The distribution of radiation over the image plane, due to a quasimonochromatic spatially coherent distribution of radiating sources in the object plane restricted to the isoplanatic region is (Dumontet, 1955)

$$g(x_i,y_i) = \iint_{-\infty}^{\infty} o(x_o,y_o)\ \bar{k}(x_o-x_i,y_o-y_i)dx_o dy_o \tag{1.15}$$

Note that this has the form of a convolution integral, which may be conveniently written in shorthand form after a scale change of eqn (1.3) is allowed so that all the coordinates are the same.

$$g(x_i,y_i) = p(x_i,y_i) \otimes \bar{k}(x_i,y_i) \tag{1.16}$$

where $\otimes$ denotes convolution.

Generally, the final image of an optical system is photographed or evaluated by eye. In either case, it is the image intensity that is measured.

Thus, to enable images from different optical systems to be compared, the amplitude distribution in the image plane is converted to an intensity distribution in the image plane.

$$\begin{aligned} i(x_i,y_i) &= |g(x_i,y_i)|^2 \\ &= \left|\iint_{-\infty}^{\infty} o(x_o,y_o)\ \bar{k}(x_o-x_i,y_o-y_i)dx_o dy_o\right|^2 \\ &= |p(x_i,y_i) \otimes \bar{k}(x_i,y_i)|^2 \end{aligned} \tag{1.17}$$

Eqn (1.17) is the general formula for imaging a quasimonochromatic coherent object.

1.4 SPATIALLY INCOHERENT IMAGING

Consider a spatially incoherent object. As the radiating sources in the object distribution can be considered mutually independent, the corresponding elementary contributions in the image are added on an intensity basis:

$$i(x_i,y_i) = \int_0^{\infty}\iint_{-\infty}^{\infty} |o(x_o,y_o)|^2 |k(x_o,x_i,y_o,y_i,\lambda)|^2 dx_o dy_o d\lambda \tag{1.18}$$

As before, if the illumination obeys the conditions for quasimonochromatic radiation and the object distribution is within the isoplanatic region, the transmission function integral can be replaced by a convolution.

$$i(x_i,y_i) = |p(x_i,y_i)|^2 \otimes |\bar{k}(x_i,y_i)|^2 \tag{1.19}$$

Eqn (1.19) is the general formula for imaging a quasimonochromatic incoherent object distribution. Some authors have used the term "spatially incoherent monochromatic objects" but, as pointed out by Born and Wolf (1959), the term "incoherent" implies at the least quasimonochromatic radiators, since for true monochromatic radiation $\gamma(P_1,P_2,\tau)$ is unity for all P_1, P_2 and τ.

If the radiating sources in the object are partially coherent, then neither eqn (1.17) nor eqn (1.19) can be applied. There is no need to consider partially coherent sources in this thesis and many texts deal with them

rigorously (cf. Born and Wolf, 1959, Mandel and Wolf, 1961, Beran and Parrent, 1964).

1.5 DUAL PLANES

If we are looking for a certain type of information within an image or picture, a different type of presentation or organisation may enhance the required information. So, we find it useful to introduce "dual planes" complementary to the image plane.

The image plane carries the image distribution, and a dual plane carries the transformed or reorganised distribution. Useful dual planes are usually connected with the image plane by an integral transform. Transforms which frequently occur are those due to Walsh, Hadamard and Fourier. They all have peculiar advantages depending on the type of signal or information processing that is used. Dual planes are often called by the name of the transform that is used to calculate the distribution in the dual plane. We discuss the Fourier transform plane in the following sections.

There is more than one way to calculate a dual plane distribution for any object or image distribution. The most obvious way is to evaluate the integral transform analytically. This method is restricted to simple distributions.

Another way to calculate a dual plane distribution is to use a digital computer. The object distribution is sampled and the computer program produces a sampled representation of the dual plane distribution.

A third way to calculate a dual plane distribution is to use an analogue device. The operation of some physical devices is described by an integral transform. Since the integral transform describes the device, the device may be used to represent the integral transform. Consequently, a particular device may be operated as a computer to "calculate" a dual plane distribution. An obvious example is using a lens, illuminated by coherent light, to compute a Fourier transform (FT) distribution. Many integral transforms do not have a convenient physical counterpart and either an analogue or a digital computer must be used to calculate the distribution.

1.5.1 The Fourier Transform Plane

If in an object or image plane there is a distribution $g(x,y)$, then there is in the FT plane a distribution $G(\alpha,\beta)$ and these two distributions are connected by the two-dimensional FT's:

$$G(\alpha,\beta) = \iint_{-\infty}^{\infty} g(x,y)\ e^{-j2\pi(\alpha x+\beta y)}\,dxdy \tag{1.20}$$

$$g(x,y) = \iint_{-\infty}^{\infty} G(\alpha,\beta)\ e^{j2\pi(\alpha x+\beta y)}\,d\alpha d\beta \tag{1.21}$$

Often the FT operation is denoted by a double ended arrow connecting the two distributions:

$$g(x,y) \longleftrightarrow G(\alpha,\beta) \tag{1.22}$$

The distribution, $G(\alpha,\beta)$, in the FT plane may be called the spatial frequency distribution, or spectrum of the object $g(x,y)$.

Another way to obtain a Fourier transform distribution of g(x,y) is to calculate it using a digital computer. Firstly, the continuous distribution g(x,y) is sampled into an N x N matrix

$$g_s(k,\ell) = \sum_{k=1}^{N} \sum_{\ell=1}^{N} g(x,y)\ \delta(k\Delta x - x)\ \delta(\ell\Delta y - y) \qquad (1.23)$$

where k,ℓ are integers

$g_s(k,\ell)$ is the sampled distribution

$\delta(\ \ )$ is the Dirac delta function

and Δx, Δy are the increment intervals along the x and y axis. Necessary conditions for an accurate representation of g(x,y) are that Δx, Δy are smaller than the minimum resolvable detail in g(x,y) and that

$$g(x,y) = 0 \qquad N\Delta x < x < 1 \quad \text{and} \quad N\Delta y < y < 1 \qquad (1.24)$$

The sampled Fourier transform distribution is calculated in a computer using the 'Fast Fourier Transform' (Bergland, 1969)

$$G_s(m,n) = \frac{1}{N^2} \sum_{k=1}^{N} \sum_{\ell=1}^{N} g_s(k,\ell)\ e^{\frac{j2\pi}{N^2}(km+\ell n)} \qquad (1.25)$$

where m and n are integers.

The continuous Fourier transform distribution is calculated from $G_s(n,m)$:

$$G(\alpha,\beta) = \sum_{m=1}^{N} \sum_{n=1}^{N} G_s(m,n)\ \frac{\text{Sin}\ \pi(x-m\alpha)}{(x-m\alpha)} \cdot \frac{\text{Sin}\ \pi(y-n\beta)}{(y-n\beta)} \qquad (1.26)$$

A major problem in calculating the FT distribution digitally is that, if $g(x,y)$ is restricted in space, then, $G(\alpha,\beta)$ exists over an infinite α,β plane. This means that many samples of $G(\alpha,\beta)$ must be taken, i.e. N must be very large.

A third way to calculate $G(\alpha,\beta)$ from $g(x,y)$ is an analogue method, using a glass lens. Consider an optical system similar to Fig. 1.1. A coherent wavefront illuminates a transparency in the object plane of the lens so that the amplitude of the wavefront behind the transparency represents $g(x,y)$. The distribution of light in the rear focal plane $U(x_f,y_f)$ is given by Goodman (1968) as

$$U(x_f,y_f) = e^{\frac{j2\pi}{2\bar{\lambda}f}(1-\frac{z}{f})(x_f^2+y_f^2)} \iint_{-\infty}^{\infty} g(x,y)\ e^{-\frac{j2\pi}{\bar{\lambda}f}(xx_f+yy_f)}\,dxdy \tag{1.27}$$

where f is the focal length

x_f,y_f is the rear focal plane

z is the object plane distance from the lens

and $\bar{\lambda}$ is the mean wavelength.

Note that if we shift the object plane closer to the lens until z equals the focal length f, then the factor multiplying the integral becomes unity. Therefore

$$U(x_f,y_f) \propto \iint_{-\infty}^{\infty} g(x,y)\ e^{j2\pi(x\frac{x_f}{\bar{\lambda}f}+y\frac{y_f}{\bar{\lambda}f})}\,dxdy \tag{1.28}$$

and after scaling,

$$U(\alpha,\beta) \propto \iint_{-\infty}^{\infty} g(x,y)\ e^{j2\pi(\alpha x+\beta y)}\,dxdy \tag{1.29}$$

Thus, the distribution of light in the rear focal plane of the lens in Fig. 1.1 due to an object distribution $g(x,y)$ is proportional to the object's FT, which is $G(\alpha,\beta)$.

Often we are only interested in the magnitude of the FT, that is, $|G(\alpha,\beta)|$. Referring back to eqn (1.27) and taking the magnitude of $U(x_f,y_f)$ we get

$$|U(x_f,y_f)| = \left|\iint_{-\infty}^{\infty} g(x,y)\ e^{-\frac{j2\pi}{\lambda f}(xx_f+yy_f)}\ dxdy\right| \qquad (1.30)$$

As eqn (1.30) is independent of the object to lens distance z, the object transparency representing the distribution $g(x,y)$ may be placed anywhere in the lens system to the left of the rear focal plane and the distribution of light in the rear focal plane is proportional to $|G(\alpha,\beta)|$.

The accuracy of the FT distribution $G(\alpha,\beta)$, usually depends on how accurately the input transparency represents the desired object distribution $g(x,y)$. If $g(x,y)$ is a complex amplitude distribution, then it is not easy to produce an accurate transparency. A method of representing a complex distribution by a purely positive distribution is described in Chapter 2 and is called holography. Holographic techniques are also required to measure the complex FT distribution $U(\alpha,\beta)$ that exists in the rear focal plane of the lens.

1.6 APPLICATIONS OF THE FOURIER TRANSFORM

We show how the concept of a FT dual plane can be used to analyse the performance of a well corrected optical system and how it can be used to improve an image produced by a poor optical system.

1.6.1 Diffraction Limited Imaging

Consider a multielement lens that has aberrations and has a limited diameter, used to image a coherently radiating object. The operation of this lens is described by eqn (1.16) and dropping the subscripts

$$g(x,y) = p(x,y) \otimes \bar{k}(x,y) \tag{1.31}$$

It is convenient to divide the aberrations of a lens into two parts. One part of the total aberrations is due to the lens having a limited diameter and the other part is due to all the remaining aberrations. These are the Seidal aberrations and aberrations due to non-uniform media between the object plane and image plane. Rewrite eqn (1.31) to give

$$g(x,y) = p(x,y) \otimes a(x,y) \otimes h(x,y) \tag{1.32}$$

where $a(x,y)$ is the amplitude psf due to the limited diameter of the lens and $h(x,y)$ is the amplitude psf due to the remaining aberrations.

Let us look at the effect on the image of a lens having a limited diameter. Define

$$b(x,y) \equiv p(x,y) \otimes a(x,y) \tag{1.33}$$

The FT equivalent of eqn (1.33) is

$$B(\alpha,\beta) = P(\alpha,\beta)\, A(\alpha,\beta) \tag{1.34}$$

$A(\alpha,\beta)$ is known as the coherent transfer function of a lens which has a limited diameter but no other aberrations, and it represents the weighting that is given to the spatial frequencies in the image spectrum $P(\alpha,\beta)$. Thus the coherent

transfer function that produces a perfect image is unity for all α,β. This means that $a(x,y)$ is a 'delta' function where a 'delta' function is defined as infinitely large at its origin and zero everywhere else. To find the effect of the lens on any image it is sufficient to determine the coherent transform function of the lens. Goodman (1968) shows that the coherent transfer function of a limited-diameter lens is

$$\begin{aligned} A(\alpha,\beta) &= 1 \qquad \alpha^2+\beta^2 \leqslant (\zeta/2\bar{\lambda})^2 \\ &= 0 \qquad \alpha^2+\beta^2 > (\zeta/2\bar{\lambda})^2 \end{aligned} \tag{1.35}$$

where ζ is the diameter of the lens.

Substituting eqn (1.35) into eqn (1.34) ensures that $B(\alpha,\beta)$ has no spatial frequencies above $(\zeta/2\bar{\lambda})$. The only way that $B(\alpha,\beta)$ can contain all the spatial frequencies of any image produced by a perfect lens is if the diameter of the lens ζ is infinitely large. As all lens are limited in diameter, no physical lens can produce the perfect image $p(x,y)$. We now introduce the concept of "the best available image" which is the best image a lens with a limited diameter can produce, provided that it has no other aberrations. So for a coherent optical system, $b(x,y)$, defined by eqn (1.33), is the best available image. Thus, a lens with aberrations that has a psf $h(x,y)$, produces an image amplitude distribution $g(x,y)$ given

$$g(x,y) = b(x,y) \otimes h(x,y) \tag{1.36}$$

and the FT equivalent of eqn (1.36) is

$$G(\alpha,\beta) = B(\alpha,\beta)\, H(\alpha,\beta) \tag{1.37}$$

Now consider this same lens used to image a quasi-monochromatic incoherent object. Using the same reasoning as applied to the lens when used with coherent radiation, eqn (1.19) may be rewritten

$$i(x,y) = |p(x,y)|^2 \otimes |a(x,y)|^2 \otimes |h(x,y)|^2 \qquad (1.38)$$

Usually it is convenient to define

$$c(x,y) \equiv |p(x,y)|^2 \otimes |a(x,y)|^2 \qquad (1.39)$$

$$d(x,y) \equiv |h(x,y)|^2 \qquad (1.40)$$

and so $c(x,y)$ is the best available image intensity that an incoherent limited-diameter imaging system can produce, and $d(x,y)$ is the intensity of the psf. (Often authors refer to $d(x,y)$ as "the psf" especially when they are dealing with completely incoherent systems.)

The FT of eqn (1.39) is

$$C(\alpha,\beta) = [P(\alpha,\beta) * P^*(\alpha,\beta)][A(\alpha,\beta) * A^*(\alpha,\beta)] \qquad (1.41)$$

where $*$ denotes correlation.

From eqn (1.35), we write the transfer function of eqn (1.41) as

$$A(\alpha,\beta) * A^*(\alpha,\beta) = 1 - \bar{\lambda}/\zeta(\alpha^2+\beta^2)^{\frac{1}{2}} \qquad \alpha^2+\beta^2 < (\zeta/\bar{\lambda})^2$$

$$= 0 \qquad \alpha^2+\beta^2 \geq (\zeta/\bar{\lambda})^2 \qquad (1.42)$$

Substituting eqn (1.42) into eqn (1.41) ensures that $C(\alpha,\beta)$ has no spatial frequencies above $\zeta/\bar{\lambda}$.

It is usual to call a lens diffraction limited if the lens has no measurable aberrations other than that due to the lens having a limited diameter.

1.6.2 Enhancing an Image Contaminated by Noise

Consider an incoherent image formed by a diffraction limited optical system which also has added an unknown random noise intensity distribution over the x,y plane of $n(x,y)$. The image formed in the absence of noise is the best available image $c(x,y)$. The 'noisy' image is $i(x,y)$, given by

$$i(x,y) = c(x,y) + n(x,y) \qquad (1.43)$$

We want to obtain a close approximation to $c(x,y)$ from $i(x,y)$.

Writing eqn (1.43) in terms of its FTs,

$$I(\alpha,\beta) = C(\alpha,\beta) + N(\alpha,\beta) \qquad (1.44)$$

Because the lens aperture is finite and has a diameter of 'ζ' units, the spatial frequencies of the best available image $c(x,y)$ do not extend beyond $\zeta/\bar{\lambda}$. This is given by eqn (1.42) in section 1.3.1 . However, the spatial frequencies of the noise spectrum $N(\alpha,\beta)$ usually extend beyond $\zeta/\bar{\lambda}$. If we filter out all the spatial frequencies of $I(\alpha,\beta)$ that exist beyond $\zeta/\bar{\lambda}$, then we are reducing the noise content while leaving the best available image unaffected. Let the filtered spatial frequency distribution be denoted by a superscript +, thus

$$\begin{aligned} I^{+}(\alpha,\beta) &= C(\alpha,\beta) + N(\alpha,\beta) && \alpha^2+\beta^2 \leqslant (\zeta/\bar{\lambda})^2 \\ &= 0 && \alpha^2+\beta^2 > (\zeta/\bar{\lambda})^2 \end{aligned} \qquad (1.45)$$

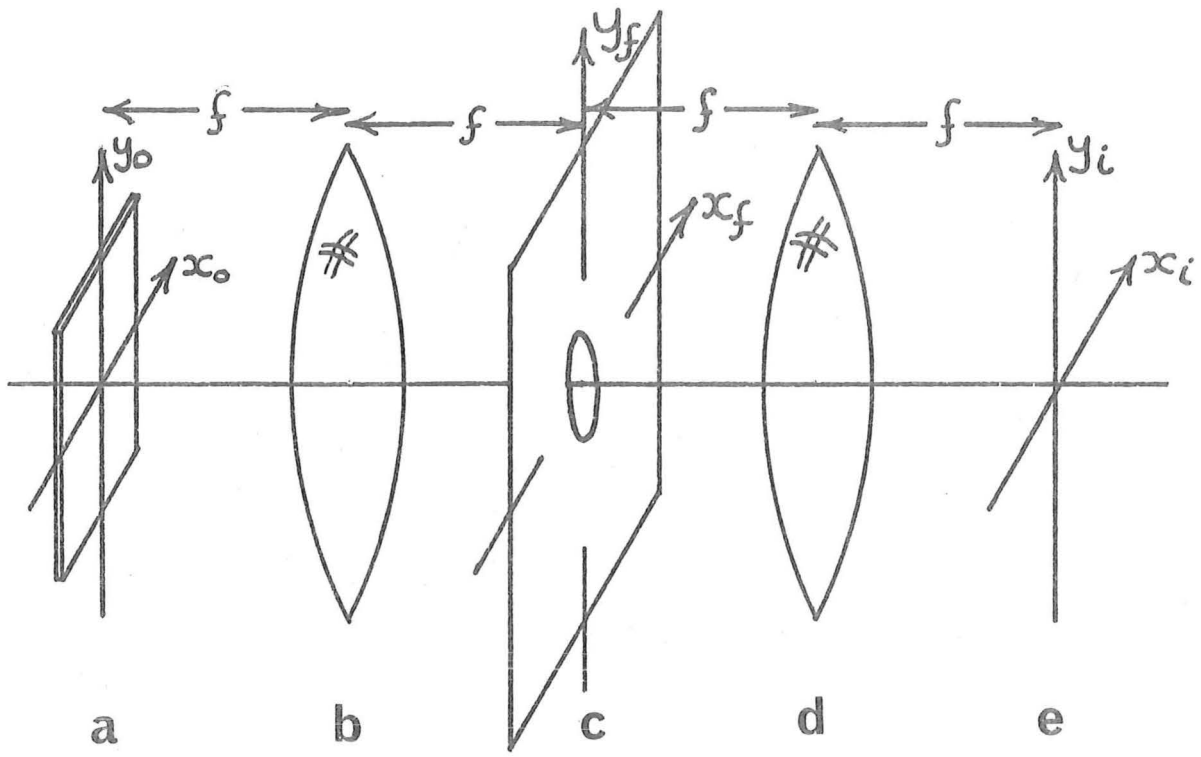


Fig. 1.2 Spatial Filtering System

a) Object plane b) First FT lens

c) Filter plane d) Second FT lens

e) Image plane.

The light distribution behind the filter is $I^+(\alpha,\beta)$. This light is passed through the second lens in Fig. 1.2 and a filtered image is produced at the rear focal plane, which is the image plane.

Consider this noise-reducing-process carried out digitally. The photograph that represents the noise image $i(x,y)$, is placed in a scanning system, so that eventually a sampled digital approximation to $i(x,y)$, is stored in the computer. The FT of $i(x,y)$ is calculated and this is a sampled digital approximation to $I(\alpha,\beta)$ which is then filtered to produce $I^+(\alpha,\beta)$. In general, a digital filter has much greater flexibility than an optical filter, since non-linear filtering operations may be performed.

The FT of $I^+(\alpha,\beta)$ is computed and this represents the spectrum of the filtered image $i^+(x,y)$. This sampled digital approximation to $i^+(x,y)$ is plotted or printed out to give the filtered image.

1.8 IMAGE PROCESSING

Generally the image that is to be processed is the output of some distorted imaging system. If this system is a coherent imaging system, it produces an image $g(x,y)$ determined by,

$$g(x,y) = b(x,y) \otimes h(x,y) \tag{1.47}$$

where $b(x,y)$ is the amplitude distribution of the best available image, and $h(x,y)$ is the amplitude of the psf.

If the system is an incoherent imaging system (as most are), then from section 1.4 the image intensity $i(x,y)$ is given by

$$i(x,y) = c(x,y) \otimes d(x,y) \tag{1.48}$$

where $c(x,y)$ is the intensity distribution of the best available image and $d(x,y)$ is the intensity of the psf.

Since image processing techniques are independent of the type of system that formed the distorted image, we show processing for only a coherent imaging system. We write eqn (1.47) in terms of its FT

$$G(\alpha,\beta) = B(\alpha,\beta)\, H(\alpha,\beta) \tag{1.49}$$

To recover $B(\alpha,\beta)$ from $G(\alpha,\beta)$, the spectrum of the

distorted image $G(\alpha, \beta)$ is multiplied by the inverse filter function $1/H(\alpha,\beta)$. The inverse filter may be very difficult to produce, since if $H(\alpha,\beta)$ is zero anywhere in the α,β plane, $1/H(\alpha,\beta)$ has an infinite transmissivity there. Denote the best approximation to the inverse filter by $\overline{(1/H(\alpha,\beta))}$. The processed image spectrum produced by this filter is $\hat{B}(\alpha,\beta)$ given by

$$\hat{B}(\alpha,\beta) = B(\alpha,\beta)\, H(\alpha,\beta)\, \overline{(1/H(\alpha,\beta))} \tag{1.50}$$

If the processing is performed digitally, $\overline{(1/H(\alpha,\beta))}$ is a discrete distribution of complex numbers and the magnitude of any one sample may lie between zero and the largest number that the computer can handle.

If the processing is performed optically $\overline{(1/H(\alpha,\beta))}$ is a continuous distribution with a limited range in magnitude, since it is not possible to make an optical filter that has a transmissivity greater than unity. Also if $1/H(\alpha,\beta)$ is a complex distribution then the optical filter must be produced by a holographic method or by using special phase retarding plates.

The advantages of coherent optical processing are a very large number of resolvable elements and a rapid speed of processing. The number of resolvable elements in a coherent optical system is limited by the grain size of the film used to represent the object distribution and to record the processed image. In much of the work reported here the film used was Kodak 649.F emulsion, either on glass plate or on a flexible acetate backing. This film has a resolution of up to 5000 lines/mm. Thus, there are up to 25×10^8 resolvable elements in every square centimeter of film.

Another advantage is the processing speed. Once the optical system has been set up and aligned, the processing speed is limited by the time taken to change the object transparencies and to record the processed image. Optical processing is virtually instantaneous. However, it may take a considerable time to set the optical system up so that aberrations due to misalignments are minimised.

The main disadvantages of coherent optical processing are noise due to dust, and the lack of dynamic range in the film. The dust in the atmosphere settles on the lens surfaces and gives unwanted diffraction effects. This can be minimised by suitable airlocks and extraction fans in the optical laboratory. The lack of dynamic range in the film has a greater effect. The smallest change in transmissivity that can be measured is approximately .05, while the maximum and minimum transmissivities are 0.9 and 0.05. Thus the dynamic range is approximately 12 db.

The advantages of a digital processing system are a noise-free environment, a large dynamic range and the ability to compute non-linear operations. Although round-off error may be considered as noise in a digital system, it is small and may often be ignored. The dynamic range of any sample in a digital system is determined by the programs used. In the IBM 360/44 computer, the largest integer number used is 32,000, so that the dynamic range is 45 db. It is not always necessary to calculate an image to this degree of accuracy, but many of the commercial subroutines used require this order of accuracy to prevent errors accumulating during a computation. The greatest advantage in a digital system is the

flexibility in being able to compute non-linear operations (Oppenheim et al., 1968). Perhaps the easiest way to illustrate this flexibility is to consider an imaging system contaminated with multiplicative noise (cf. with section 1.6.2) given by

$$i(x,y) = c(x,y)\, n(x,y) \tag{1.51}$$

Taking the FTs we get

$$I(\alpha,\beta) = C(\alpha,\beta) \otimes N(\alpha,\beta) \tag{1.52}$$

and it is not easy to calculate $C(\alpha,\beta)$ from $I(\alpha,\beta)$. However if the logarithm of the eqn (1.51) is taken to give

$$\log\{i(x,y)\} = \log\{c(x,y)\} + \log\{n(x,y)\} \tag{1.53}$$

and the FT of this distribution is computed, then,

$$F[\log\{i(x,y)\}] = F[\log\{c(x,y)\}] + F[\log\{n(x,y)\}] \tag{1.54}$$

Denoting the FT of a logarithm by $^{\Delta}$ and rewriting eqn (1.54) in the following form we get

$$\overset{\Delta}{I}(\alpha,\beta) = \overset{\Delta}{C}(\alpha,\beta) + \overset{\Delta}{N}(\alpha,\beta) \tag{1.55}$$

The filtering procedures that are applied in section 1.6.2 can now be applied to $\overset{\Delta}{I}(\alpha,\beta)$. After $\overset{\Delta}{I}(\alpha,\beta)$ has been filtered, the inverse FT distribution is computed and the antilog of this distribution is calculated to produce a filtered image. This image contains less noise than the image given by eqn (1.51).

To date it has not been possible to perform this type of processing optically, although a recent paper by Goodman and Kato (1974) suggests that a non-linear optical filter is possible. Non-linear optical filtering may prove to be a very useful tool for future experimenters.

The disadvantages of digital image processing are the time required to produce the filtered image, and the small number of resolvable elements in the image. There are three steps in digital processing and each step takes some time. Firstly, the transparency that represents the object distribution must be scanned and digitised. (In the rather primitive scanner available to us, this took 15 minutes.) Then the digital image is computed in the arithmetic units of the computer. The time taken for this step depends on the complexity of the calculation, and it may take up to 10 minutes to compute the processed image. The third step is to print or plot a visual impression of the processed image and this may take up to another 30 minutes depending on the type of plotter available. In the digital processor we used, the total time to produce a processed image from an object transparency varied between 20 minutes and 55 minutes.

If the object transparency is the result of an unrepeatable situation, as many space photographs are, then processing time is not important. What is more important in this type of situation is the resolution of the processed image. The quality of an image is determined by the number of resolvable elements in that image. Unfortunately, most digital processes have a very poor resolution, because of the limited size of computer memories. Typically, an image in a digital computer is stored as a matrix of 256 by 256 resolvable

elements or samples. This is only 1.5×10^5 elements and compares poorly with the 25×10^8 elements/cm^2 in a coherent optical system. These values are only indicative of the order of magnitude expected and individual systems may have more or less resolvable elements than those quoted.

1.9 REVIEW OF IMPORTANT ADVANCES IN IMAGE PROCESSING

Abbe (1873) was the first experimenter to conceive of the spatial frequency dual plane. He showed that in a spatially coherent optical system, the size of the aperture in the focal plane determines the number of resolvable points in the image plane.

Zernike (1935) utilised the spatial frequency plane (the focal plane) of a coherent optical system to develop a phase contrast microscope. Many biological specimens are almost transparent and so may be written as a phase only function.

$$o(x,y) = e^{j\phi(x,y)}$$

$$\doteqdot 1 + j\phi(x,y) \quad \text{since } \phi(x,y) << 1 \tag{1.56}$$

where $o(x,y)$ is the transmissivity of the specimen.

The intensity of its image is

$$i(x,y) \doteqdot 1 + \phi^2(x,y)$$

$$\doteqdot 1 \tag{1.57}$$

Zernike introduced a phase retarding plate in the focal plane, so that it retarded the central or DC term of the spatial

frequency plane by π radians. Therefore

$$i(x,y) = |e^{j\pi/2} + j\phi(x,y)|^2$$

$$= 1 + 2\phi(x,y) \qquad (1.58)$$

So now, the image intensity is linearly related to the phase introduced by the specimen. This type of deliberate distortion to enable an image to be seen more clearly is termed image enhancement.

Marchael et al. (1953) were the first experimenters to use an amplitude and phase filter. They took a badly focused image, and attempted to correct the effects of misfocusing to 'restore' the image. The effects of a misfocused coherent imaging system can be written as a convolution.

$$g(x,y) = b(x,y) \otimes h(x,y) \qquad (1.59)$$

where $g(x,y)$ is the misfocused image

$b(x,y)$ is the best available image

$h(x,y)$ is the psf of the misfocused system.

Equally

$$G(\alpha,\beta) = B(\alpha,\beta)\ H(\alpha,\beta) \qquad (1.60)$$

They first constructed an intensity-only filter $1/|H(\alpha,\beta)|^2$, and then etched a phase-only filter in glass, which was stuck to the intensity filter so that both filters approximated $1/H(\alpha,\beta)$. This composite filter was then multiplied to $G(\alpha,\beta)$, and the subsequent FT gave an approximation to the correct image $b(x,y)$. This type of

correction is generally termed image restoration.

Holography is probably the most revolutionary idea to emerge in optics during the last thirty years. Developed by Gabor (1948,49) in an attempt to improve the imaging ability of electron microscopes, holograms now have a wide variety of practical uses. He proposed to 'freeze' the wavefront emanating from an object. Then, this 'object' wavefront could be reconstructed later. Since a radiating wavefront has both amplitude and phase, and no known medium could preserve the phase, it seemed impossible. However, he proposed that the object wavefront be interfered with a known and easily duplicated 'reference' wavefront, and then, an intensity-only recording be made of the interference fringes of the two fields. This interference pattern he called a hologram. If this hologram is illuminated by a similar wavefront to the 'reference', the desired object wavefront is reconstructed in both amplitude and phase.

A more thorough analysis of holography is given in Chapter 2.

One of the problems with Gabor's original hologram was that the reconstructed wavefront was always mixed with some unwanted wavefronts. This was owing to the limited temporal coherence of the light sources available in the 1940's. The most coherent source at that time was a filtered arc source. An incandescent arc is struck in an atmosphere of inert gases and the spectrum of the radiated light is composed of a number of narrow bandwidth (approximately 80 $\mathring{A}$) spectral "lines". This light is passed through a bandpass filter which attenuates all but one spectral "line". The coherence

distance along any ray path of this light is less than 1 mm. The small coherence distance limited the application of holography until the laser was invented which increased the coherence distance to several centimetres. Leith and Upatniks (1963, 1964) were able to introduce the reference beam at an angle to the optical axis which meant that on reconstruction the object wavefront could be viewed separately from the unwanted wavefronts. This gives a clear image of the original object.

Applying this off axis technique to image processing, Vander Lugt (1964) was able to make amplitude and phase filters in one step (cf. Marchael et al., 1953). Generally he made filters of the type $H(\alpha,\beta)$ and used matched filtering techniques to identify shapes. Selecting key words in texts has been its main use to date.

Digital computers had now reached a stage where they could be used in image processing. Brown and Lohman (1966) developed a scheme by which the output of a computer plotter could simulate $H(\alpha,\beta)$ and $1/H(\alpha,\beta)$. They then photoreduced the plotter output onto a transparency and used this as a filter in the spatial frequency plane of a coherent optical system. Their technique was similar in concept to the wavefront reconstruction proposed by Gabor, so the filters were called computer generated holograms. Many different computer generated holography schemes have been advanced, and these are explained in Chapter 2.

Purely digital image restoration and image enhancement schemes have been advanced by H.C. Andrews (1970) and Oppenheim et al. (1968). Most digital image processing schemes

were limited by the core storage of the computers available, and by the time taken to perform the 2-dimensional FT. The first limitation is gradually being reduced as computer memories get larger, and the second limitation was greatly reduced by the development of the Fast Fourier Transform (FFT) popularised by Cooley and Tukey (1965)

Similar fast transforms now exist for the Hadamard transform (Pratt et al. 1969). The biggest advantage of digital spatial filtering is that non-linear transfer functions can be tried easily (Oppenheim et al., 1968), especially in the case of image enhancement where it can be largely a case of trying a function and then looking at the result. Andrews (1970) has been using this type of technique on digitised space pictures.

Another field which is being investigated is bandwidth reduction of an image which involves determining how much information can be left out of an image and yet still leave a 'good' image (Anderson and Huang, 1971).

Recently Stroke et al. (1972) have been the leading proponents of purely optical image processing. They have developed a two part process to construct a filter of the type $1/H(\alpha,\beta)$ quite accurately. They have used such a filter to improve the resolution of an electron microscope down to 2.5 Å. This type of filtering has also been used to correct motion blurred photographs, and to correct for misfocusing of cameras (NASA document SP193 , 1969). Our experience is that filters constructed by Stroke's method are not easy to make and require very carefully controlled conditions to work at all.

Not all the various types of processing have been mentioned and the reader's attention is directed to Huang et al. (1971) and Oppenheim et al. (1968) where the field is reviewed very thoroughly.

CHAPTER 2

HOLOGRAPHY AND COMPUTER GENERATED HOLOGRAPHY

This chapter is divided into two parts. The first outlines the principles of holography and explains the construction of most types of holograms. The second part deals with a synthetic hologram where the hologram is calculated in a digital computer and the object wavefront is reconstructed optically. There is a short review of the main types of computer generated holograms.

2.1 GENERAL PRINCIPLES OF HOLOGRAPHY

Holography or wavefront reconstruction is an attempt to preserve the phase of a radiating wavefront. If the wavefront is incident on a recording medium, what is required is that the medium records both the phase and the amplitude of the incident wavefront. All known forms of recording medium require a considerable amount of energy to activate them. The time taken to record a wavefront may be many times the wavefront's time period. Thus, unless special precautions are taken, the phase information is lost. Photographic film for instance records only the intensity of the incident signal. Let the incident object wavefront be $O(\alpha,\beta)$. The intensity is given by

$$I(\alpha,\beta) = |O(\alpha,\beta)|^2$$

$$= O(\alpha,\beta)\, O^*(\alpha,\beta) \qquad (2.1)$$

If this intensity is recorded on a sheet of film, then all information about the phase is lost. (The linearity of the recording of the intensity depends on the type of film and developer used, and this is explained more fully in Appendix A.)

If a reference wavefront $R(\alpha,\beta)$ is now added to the object wavefront $O(\alpha,\beta)$, their combined intensity is $I(\alpha,\beta)$ where

$$\begin{aligned} I(\alpha,\beta) &= |O(\alpha,\beta) + R(\alpha,\beta)|^2 \\ &= |R(\alpha,\beta)|^2 + |O(\alpha,\beta)|^2 + R^*(\alpha,\beta)\, O(\alpha,\beta) \\ &\quad + R(\alpha,\beta)\, O^*(\alpha,\beta) \end{aligned} \tag{2.2}$$

The third term of the above expansion contains both the amplitude and phase of the object wavefront $O(\alpha,\beta)$. The exposure of the film is defined as the intensity of the illumination multiplied by the exposure time ΔT. The exposure and the development of the film used to record $I(\alpha,\beta)$ are set so that the amplitude transmissivity of the developed film has a mean value of 0.5. Appendix B explains why this is so. Since the third and fourth terms of eqn (2.2) have average values of zero, the exposure of the film is set by the average value of $|R(\alpha,\beta)|^2$ and $|O(\alpha,\beta)|^2$ multiplied by the exposure time ΔT. If the amplitude transmissivity across the developed film is $t(\alpha,\beta)$ (which must lie between 0 and 1), then

$$\begin{aligned} t(\alpha,\beta) &= |R(\alpha,\beta)|^2\Delta T + |O(\alpha,\beta)|^2\Delta T + R^*(\alpha,\beta)O(\alpha,\beta)\Delta T \\ &\quad + R(\alpha,\beta)O^*(\alpha,\beta)\Delta T \end{aligned} \tag{2.3}$$

If the transparency representing $t(\alpha,\beta)$ by a wavefront is illuminated by a wavefront identical to the reference wavefront, the complex field $U(\alpha,\beta)$ just behind the transparency is

$$\begin{aligned} U(\alpha,\beta) &= R(\alpha,\beta)\, t(\alpha,\beta) \\ &= R(\alpha,\beta)|R(\alpha,\beta)|^2\Delta T + R(\alpha,\beta)|O(\alpha,\beta)|^2\Delta T \\ &\quad + |R(\alpha,\beta)|^2 O(\alpha,\beta)\Delta T + R^2(\alpha,\beta)O^*(\alpha,\beta)\Delta T \end{aligned} \tag{2.4}$$

In most applications $R(\alpha,\beta)$ is a plane wave with a constant amplitude and phase over the α,β plane and it is often convenient to normalise the amplitude of the wavefronts so that

$$R(\alpha,\beta) = 1 \tag{2.5}$$

for all α,β. Substituting eqn (2.5) into eqn (2.4) gives

$$\hat{U}(\alpha,\beta) = \Delta T + |O(\alpha,\beta)|^2\Delta T + O(\alpha,\beta)\Delta T + O^*(\alpha,\beta)\Delta T \tag{2.6}$$

where $\hat{U}(\alpha,\beta)$ refers to the normalised amplitudes. There are four wavefronts in $\hat{U}(\alpha,\beta)$, since there is a separate wavefront for each term in eqn (2.6). The first wavefront emanating from behind the transparency is a plane wave. The second wavefront is proportional to $|O(\alpha,\beta)|^2$. The third wavefront is proportional to $O(\alpha,\beta)$, so that both the amplitude and the phase of the object wavefront have been reconstructed from an intensity-only measurement. A transparency used in this way is called a hologram. More generally, it may be considered that a hologram is any intensity measurement that records the phase and the amplitude of a radiating wavefront. The fourth wavefront emanating from just behind the transparency

(hologram) is proportional to $O^*(\alpha,\beta)$. Both $O^*(\alpha,\beta)$ and $O(\alpha,\beta)$ form images in space and the two images are usually separated from each other. The type of image that is formed by the two wavefronts may be real or virtual and the particular way in which the two images are separated from the unwanted wavefronts gives rise to the many different types of holograms.

2.2 EXPERIMENTAL HOLOGRAPHY

Although most holograms are constructed and the wavefronts reconstructed at optical wavelengths, some holograms are constructed at one wavelength and the wavefront reconstructed at another. The following section describes holograms formed by various different kinds of coherent radiation.

2.2.1 Optical Holograms

Gabor (1948, 1949) is the originator of holography. In attempting to improve the resolution of an electron microscope he formulated the holographic principle of combining a known reference wavefront with an object wavefront and then recording only their combined intensity. For the two wavefronts to combine, or interfere, they must be coherent or at least partially coherent. In most of Gabor's experiments, he used a filtered arc source radiating through a pinhole as this was the most efficient way to obtain quasimonochromatic coherent illumination at that time. However, a filtered arc source gives illumination that has very little temporal coherence, and for interference to occur the difference in path lengths between the reference and the object wavefronts from the source to the hologram plane must be less than 1 mm. This

requirement places strict limits on the layout of the optical components. Fig. 2.1 illustrates the geometry of an optical system capable of producing a Gabor hologram. The disadvantage of a Gabor hologram is that it is not possible to completely separate the reconstructed object wavefront from the other three wavefronts. This is illustrated in Fig. 2.1(b).

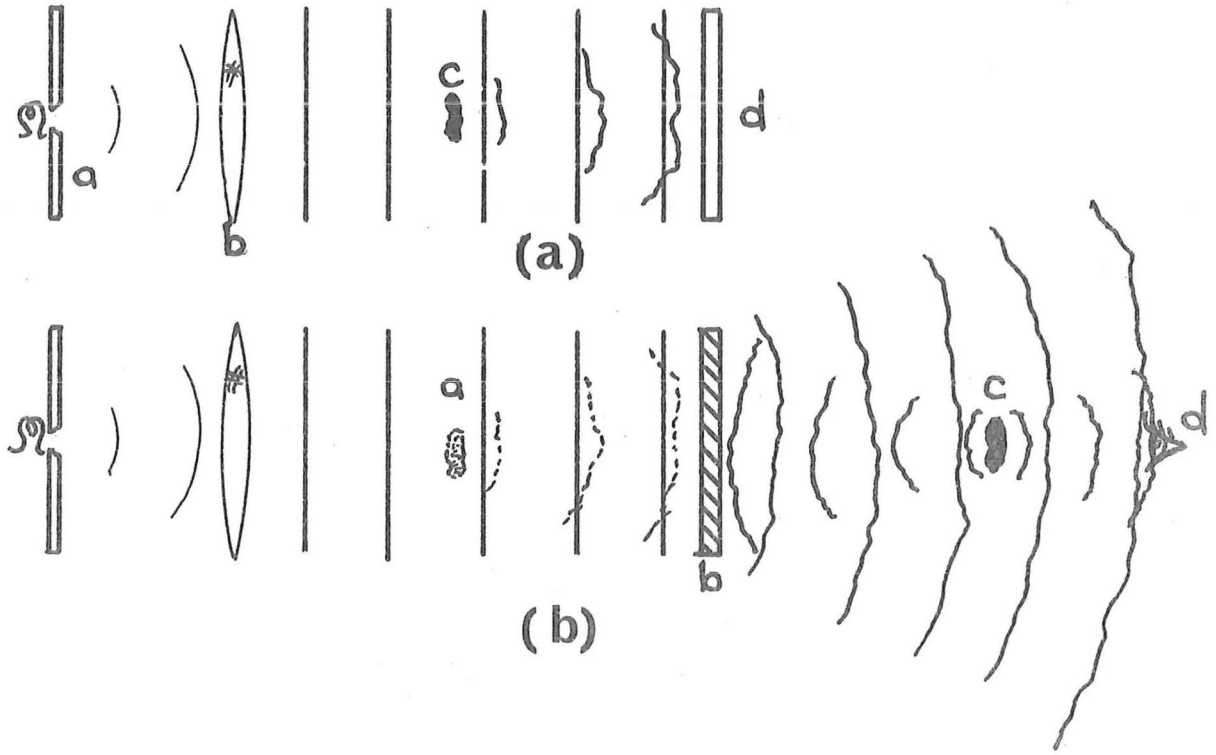


Fig. 2.1 Optical Geometry for a Gabor Hologram

(a) Constructing the hologram: a) the arc source and pinhole; b) collimating lens; c) object; d) film.

(b) Reconstructing the object wavefront: a) virtual image; b) hologram; c) real image; d) viewing area for real and virtual images.

To lessen the visual confusion, the magnitude of the object wavefront is made much less than the magnitude of the reference wavefront. This reduces eqn (2.6) to the first,

third and fourth terms, since $|O(\alpha,\beta)|^2$ becomes negligible. The object wavefront is made weaker than the reference wavefront by ensuring that the object transparency is highly transmissive. That is, most of the illuminating spatially coherent waves from the source pass through the object transparency unaffected. When these waves arrive at the hologram plane they form the reference wavefront. A small proportion of the illumination is diffracted by the object. These diffracted waves, when they arrive at the hologram plane, form the object wavefront $O(\alpha,\beta)$. This is the Fresnel diffraction pattern of the object. As the magnitude of $O(\alpha,\beta)$ is much less than $R(\alpha,\beta)$ eqn (2.6) is rewritten, where R has again been normalised to unity to produce

$$\hat{U}(\alpha,\beta) = \Delta T + O(\alpha,\beta)\Delta T + O^*(\alpha,\beta)\Delta T \qquad (2.7)$$

Even now there is still some visual confusion between the images formed by $O(\alpha,\beta)$ and $O^*(\alpha,\beta)$ when the wavefronts are reconstructed from the hologram. This is because the two images are in-line and it is necessary to look through one image to see the other as in Fig. 2.1(b). The visual confusion arising from the two images limited the application of holography when it was originally proposed by Gabor.

When the laser was developed it provided a quasi-monochromatic source whose temporal coherence extended over many thousands of wave periods. Consequently, the coherence distance of the laser is several centimeters which means that the path lengths of the reference wavefront and the object wavefront from the pinhole to the hologram plane now may be different by several centimetres. The hologram plane is no longer required to be perpendicular to the reference wavefront

and object wavefronts. Leith and Upatniks (1964) introduced the reference wavefront at an angle to the hologram plane. On reconstructing the wavefronts from the developed hologram, the two conjugate images are reconstructed at different angles to each other. This is illustrated in Fig. 2.2.

This type of hologram may also be produced when the Fresnel diffraction field is formed by reflecting wavefronts from a three-dimensional object. This is the most widely known optical hologram (Leith and Upatniks, 1964) and permits three-dimensional scenes to be reconstructed, with impressive effect.

Vander Lugt (1964) uses the off axis reference to construct matched filters for spatial filtering. The fields $O(\alpha,\beta)$ and $R(\alpha,\beta)$ are formed by taking an optical Fourier transform of two objects $o(x,y)$ and $r(x,y)$. Usually $r(x,y)$ is a point source or 'delta' function. This ensures that R has a constant amplitude with a linear phase shift across the hologram plane. The reconstructed wavefront is Fourier transformed to give two real images either side of the autocorrelation terms. The reconstructed field just behind the illuminated hologram is $U(\alpha,\beta)$, where

$$U(\alpha,\beta) \propto |R(\alpha,\beta)|^2 + |O(\alpha,\beta)|^2 + R^*(\alpha,\beta)O(\alpha,\beta) + R(\alpha,\beta)O^*(\alpha,\beta) \qquad (2.8)$$

The four separate wavefronts that constitute the reconstructed field are passed through a lens so that the FT of $U(\alpha,\beta)$ occurs at the image plane. The distribution of light in the image plane is given by

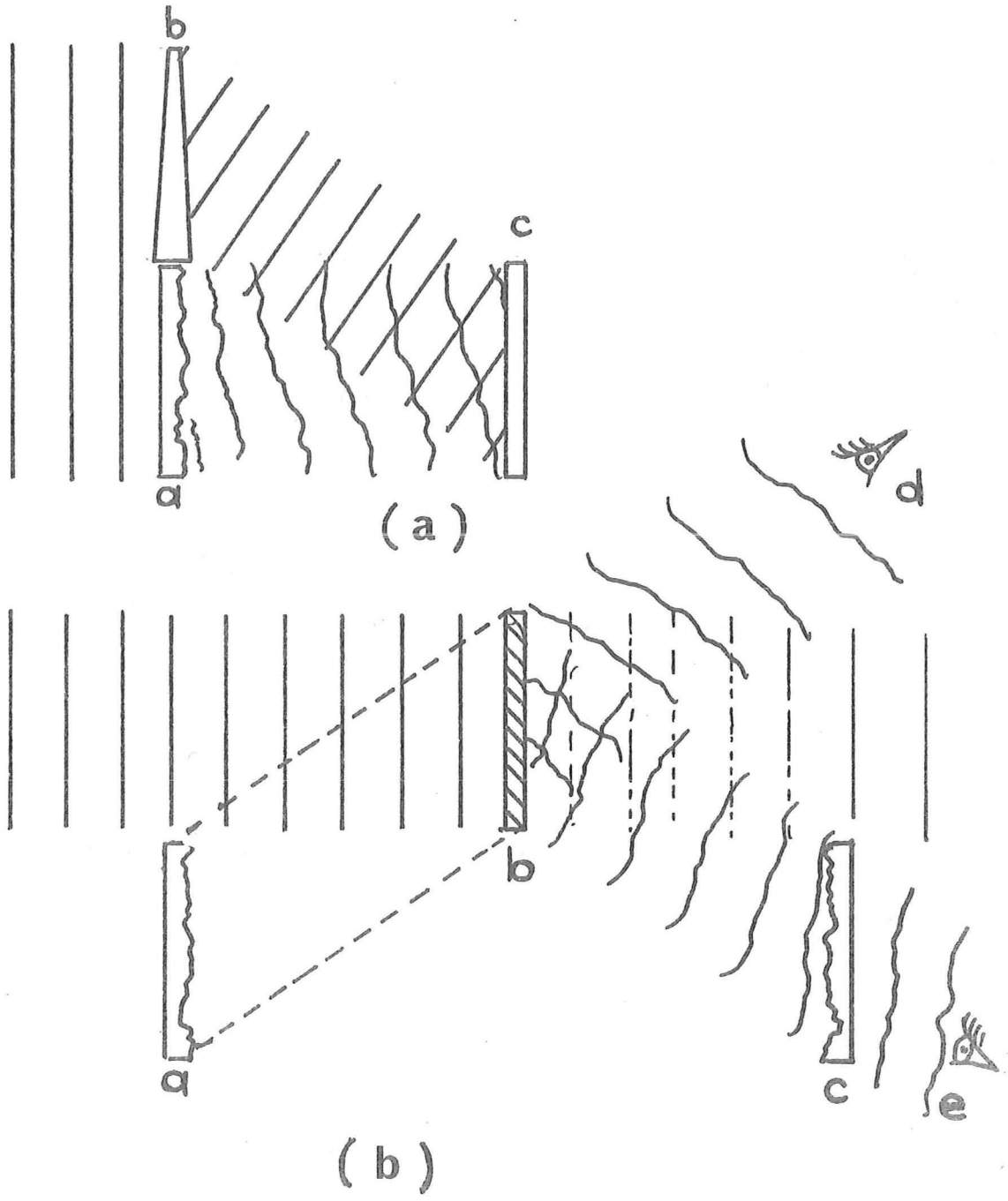


Fig. 2.2 The optical geometry for an off-axis Fresnel hologram

(a) Constructing the hologram

a) Object transparency
b) Prism
c) Unexposed film

(b) Reconstructing the object wavefront

a) Virtual image
b) Hologram
c) Real image
d) View of virtual image
e) View of real image.

$$u(x,y) \propto r(x,y) \star r(x,y) + o(x,y) \star o(x,y)$$

$$+ r(x,y) \star o(x,y) + r(x,y) \otimes o(x,y) \qquad (2.9)$$

where $\star$ denotes correlation and $\otimes$ denotes convolution. If $r(x,y)$ approximates a delta function at a distance 'a' from the optical axis, then

$$r(x,y) = \delta(x+a,y)$$

and

$$u(x,y) \propto \delta(x,y) + o(x,y) \star o(x,y) + o(x-a,y) + o(-x+a,-y) \qquad (2.10)$$

Thus, the distribution of light in the image plane has a central delta function $\delta(x,y)$, the autocorrelation of the image $o(x,y) \star o(x,y)$ and two real conjugate images displaced by $\pm a$ from the optical axis. So, if it is required that the images are to be physically separated from each other and the central terms, the original separation between $r(x,y)$ and $o(x,y)$ in the original object plane must be greater than the total extent of $o(x,y)$.

If however, $|O| << |R|$, eqn (2.6) can be approximated to

$$u(x,y) \doteqdot \delta(x,y) + o(x-a,y) + o(-x+a,-y) \qquad (2.11)$$

then the retransformed field has a central delta function and two real images displaced by $\pm$'a' units from the optical axis. Thus the separation constraints can be relaxed, so that the reference point source $r(x,y)$ is just outside $o(x,y)$.

An advantage of a FT hologram is that since both $O(\alpha,\beta)$ and $R(\alpha,\beta)$ are FTs, their combined intensity $I(\alpha,\beta)$ has constant period fringes (normally either $O(\alpha,\beta)$ or $R(\alpha,\beta)$ is a Fresnel

diffraction pattern and so the period of the fringes in their combined intensity $I(\alpha,\beta)$ varies in the same way as a Fresnel zone plate). Since the period of the fringes in a FT hologram distribution is constant over the hologram plane, the resolution of the film used to record the hologram distribution may be just sufficient to record the fringes. Film with a relatively poor resolution has been used to record FT holograms (Collier et al., 1971).

A disadvantage of a FT hologram is that film with a limited dynamic range is not always able to record both the high intensity peaks and the low intensity ripples of the hologram distribution. The dynamic range required of the film depends on the form of the object $o(x,y)$. If $o(x,y)$ is a smooth distribution then most of the energy in the FT distribution is in the central high peak at $\alpha = 0$ and $\beta = 0$. When the reference wavefront is added to the object wavefront, most of the energy in the hologram distribution is restricted to a small area around the origin (0,0), as illustrated in Fig. 2.3(a). Objects that give this type of hologram distribution are known as specular objects. If $o(x,y)$ is a diffuse object, it has rapid random phase variations across the x,y plane. However, there is little visible difference between a specular object and a diffuse object since our eyes detect only intensity. The main difference is that the FT of a diffuse object is distributed over the entire α,β plane and no one part of the film in the hologram plane has to record either a high peak of intensity or a small variation in intensity. A FT hologram of a diffuse object is illustrated in Fig. 2.3(b).

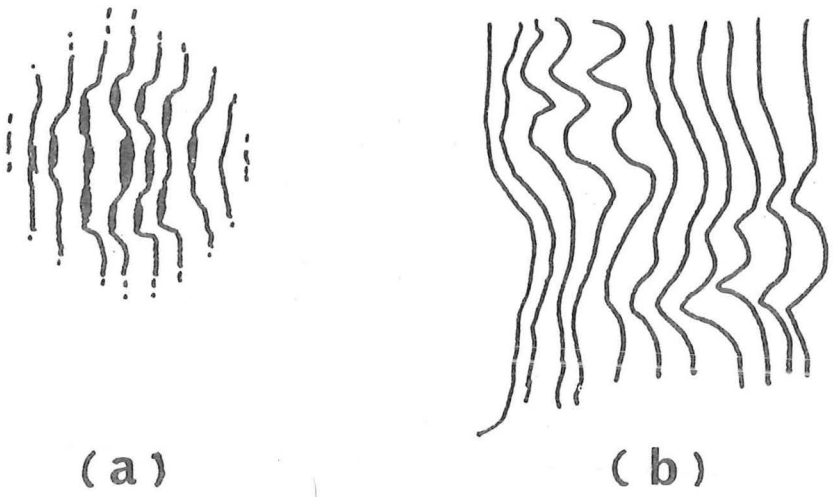


Fig. 2.3 Fringes across a FT hologram of
(a) a specular object, (b) a diffuse object.

Many optical experimenters have turned a specular object into a diffuse object by placing a fine ground glass plate just behind the transparency that represents the object. This ensures that there are small rapid phase fluctuations across the object.

Another type of constant period interference fringe hologram can be constructed without a lens. As the constant-period fringe is a characteristic of a FT hologram, this type is called a "lensless Fourier Transform" hologram. Both O and R are formed by Fresnel diffraction patterns. They reconstruct in the same way as a normal FT hologram (Develis and Reynolds, 1967).

2.2.2 Microwave Holograms

Since the mode of propagation of microwave radiaton is the same as optical radiation, the general principles of optical holography also apply to microwave holography.

Usually, a hologram is constructed by measuring the intensity of the combined reference and signal wavefronts. Then, the hologram is reconstructed at optical wavelengths to give an optical picture of the microwave source distribution of the object.

Microwave holograms have two major differences from optical holograms. The first difference is due to the vastly different wavelengths of the radiation used to construct them. The wavelengths of microwave radiation lie between 1 mm and 500 mm whereas those of visible light lie between 400 nm and 700 nm. This usually means that across any comparable radiating apertures there are many more optical wavelengths than microwave wavelengths. Consequently the radiation pattern of an optical radiator in the far field, is much more restricted (in solid angle) than the equivalent microwave radiator. As a hologram is usually constructed in a limited area of solid angle (the hologram aperture), more of the total radiated flux is received by the optical hologram.

The other difference is that image reconstruction from most microwave holograms is performed at a different frequency to the frequency that was used to construct the hologram. This means that scaling must be allowed for to correct for this frequency difference.

The usual technique used to construct a microwave hologram is to modulate a light pen that is slaved to a receiving antenna, scanning back and forth in the hologram plane. The

position and the modulation intensity of the light pen is recorded on film and the developed negative is used as an optical hologram to be reconstructed at light frequencies.

Tricoles and Rope (1967) and Checcacci et al. (1971) have reconstructed source distributions of microwave antennas at microwave frequencies and Aoki and Boivin(1971) have used a computer reconstruction technique for microwave holograms.

Microwave holograms have a major advantage over optical holograms because of the stability of microwave oscillators. The reference wavefront $R(\alpha,\beta)$ may be added electronically to $O(\alpha,\beta)$ and so $R(\alpha,\beta)$ does not need to be radiated. Cutrona et al. (1966) have applied this in 'synthetic aperture radar' and this remains the most widely used type of microwave hologram.

2.2.3 Acoustic Holograms

Acoustic fields are formed by pressure waves, but it is still possible to form an interference field between two radiated wavefronts, provided the two wavefronts are coherent.

Usually the hologram is formed at acoustic wavelengths between 1 mm and 40 mm. The hologram is then reconstructed at optical wavelengths to form an optical picture of the acoustic source distribution of the object. The object may be self-excited (as in a transmitting antenna) or used as a passive reflector to an external radiator.

As the wavelengths used by microwave and acoustic holograms are similar, their holograms have similar properties. Therefore allowances must be made for frequency scaling and radiated flux concentration for acoustic holograms in the same way that allowances are made for microwave holograms.

Several researchers have investigated holograms formed acoustically. Greene (1969) determined the source distribution of an acoustic radiator optically and Goodman (1969) reconstructed the source distribution of a radiator computationally. Accurate source distributions were calculated by Napier and Bates (1971, 1973) and some attempts have been made to form acoustic holograms directly onto photographic film (Berger, 1969).

A technique that is peculiar to acoustic holography is real time holography. This is where an instantaneous optical reconstruction is made of a continuously changing hologram. It is generally used to look into living or moving tissue, since low density tissue is opaque to optical wavelengths, but transparent to acoustic wavelengths. Usually, the object is immersed in a fluid, to lower the acoustic impedance barrier between the tissue and the source of the radiating wavefronts. The wavefronts pass through the object and disturb the surface of the fluid. At the same time, another radiating source is immersed in the fluid. This is directed at the surface of the fluid, and provided this reference source is coherent with the source radiating through the object, an interference pattern forms on the surface. This is the hologram.

The reconstruction technique is to bounce coherent visible radiation off the disturbed surface of the fluid in such a way that a visible image of the acoustic object is formed instantly.

2.3 COMPUTER GENERATED HOLOGRAMS

In many applications of holography the complex wavefront is purely artificial and is stored as a sequence of numbers in a computer. Several methods have been advanced to change an artificial wavefront into a real optical wavefront.

A simple way in which to construct a computer generated hologram is to simulate as near as possible an optical hologram. This type of computed hologram is generally called an optical analogue hologram. The major difference between the computed hologram and its optical equivalent is that the computed hologram only exists at certain grid points.

If $O(\alpha,\beta)$ is the object wavefront, then the sampled wavefront is $O_s(\alpha,\beta)$ where

$$O_s(\alpha,\beta) = \sum_{n}^{N} \sum_{m}^{N} O(\alpha,\beta)\ \delta(\alpha-n\Delta\alpha,\beta-m\Delta\beta) \qquad (2.12)$$

$O_s(\alpha,\beta)$ is stored in the computer. The reference wavefront is introduced at an angle and sampled in the same way as the object wavefront to give $R_s(\alpha,\beta)$. The sampled intensity in the hologram plane is $I_s(\alpha,\beta)$ given by

$$I_s(\alpha,\beta) = |R_s(\alpha,\beta) + O_s(\alpha,\beta)|^2 \qquad (2.13)$$

A big advantage of a computer generated hologram is that as $O_s(\alpha,\beta)$ is known completely, $|O_s(\alpha,\beta)|^2$ may be subtracted from $I_s(\alpha,\beta)$ so that the number of unwanted wavefronts emanating from behind the illuminated hologram is reduced to two. Thus the computer generated hologram can be written as

$$I_s(\alpha,\beta) = |R_s(\alpha,\beta)|^2 + R_s^*(\alpha,\beta)O_s(\alpha,\beta) + R_s(\alpha,\beta)O_s^*(\alpha,\beta) \quad (2.14)$$

Then $I_s(\alpha,\beta)$ is printed or plotted out on the computer's peripherals and photoreduced onto a transparency. This approximates the transparency that would have been recorded if $I_s(\alpha,\beta)$ had been measured purely optically and then $|O_s(\alpha,\beta)|^2$ subtracted.

Brown and Lohmann (1966) and later Lohmann and Paris (1967) developed a different technique to that used for the optical analogue hologram.

Firstly, the sampled field $O_s(\alpha,\beta)$ is considered to have amplitude and phase components at each of the sample points. Now, each sample point stored in the computer corresponds to an equivalent sample point on a sheet of white paper. The computer controls a plotter that inks in a small rectangle near the position of a sample point on the paper. The area of the rectangle is proportional to the amplitude of the object wavefront at that sample point and the distance of the rectangle from the sample point gives the phase at that sample point. Finally, when the whole object wavefront has been plotted, there are as many inked-in rectangles covering the white paper as there are sample points.

To avoid aliasing, the maximum distance that a rectangle may be drawn from its sample is such that the rectangle should not interfere with any rectangle drawn from a neighbouring sample point. A cited advantage of this technique is its binary property. Since the hologram transparency is either a transparent aperture (transmissivity of approximately 1.0) or an opaque background (transmissivity of approximately 0.0), the type of film and its development is not critical. The

film only has to record two values; the maximum density that the film can support and the minimum density which is the unexposed backing and gelatine base. Brown and Lohmann (1969) showed that it was possible to achieve some three dimensional effect and produce some grey tones.

There have been several attempts to use phase only information, mainly by Lesem et al. (1968), and Keeton (1968).

Lee (1970) uses a different method to code the synthetic object wavefront. The complex field $O(\alpha,\beta)$ is divided into real and imaginery components and each of these components is itself divided into two parts. The hologram is computed from these four non-negative functions. Coherent light is passed through the hologram so that the four functions are added together with the correct phases and the field just behind the hologram approximates $O(\alpha,\beta)$. Burckhardt (1970) shows that the synthetic object wavefront can be divided into three non-negative real numbers thus simplifying Lee's method slightly.

Haskell and Culver (1972) produce another coding scheme where they use two vertically spaced dots (usually a CRT trace) for every sample point. The vertical spacing of the dots gives the amplitude and the horizontal position of the two dots from the sample point gives the phase. This method has all the advantages of a binary hologram. This coding is probably the most suitable scheme for a CRT computer output peripheral.

Other coding schemes have been put forward by various groups but many of these rely on optical illusions to give an appearance of 3-dimensionality (King et al., 1970).

images do not overlap. (The effect of undersampling is known as aliasing in signal processing systems.) The minimum sampling period that can be tolerated is slightly shorter than the natural fringe period of the 'combined wavefronts'.

There may also be some difference in the relative sizes of the computer generated hologram and its optical equivalent. Usually the computer generated hologram is many times larger than its optical analogue, so that the images formed by the reconstructed wavefronts O_s and O_s^* are many times smaller than in the optical equivalent. A system of magnifying lenses is generally used to correct this to produce an image of conventional size.

Holography and computer generated holography now have wide applications in industry. Their uses include many types of visual displays, information storage (e.g. computer read-only memories), picture multiplexing, as well as image processing. In the following chapters we show some specific applications of holography to the processing of images.

CHAPTER 3

IMAGING THE CROSS-SECTION OF A BODY USING COMPUTER GENERATED HOLOGRAPHY

3.1 INTRODUCTION

In some biomedical experiments it is often useful to know the density variation over a particular cross-section in a body. A conventional X-ray does not give this information, since the image of one cross-section is always superimposed on the images of all the other cross-sections between the X-ray source and the photographic plate. However, it is possible to obtain the information about a particular cross-section from a series of FT operations on a number of X-ray photographs all taken at slightly differing angles.

In this chapter we outline the specific mathematical procedure to determine the cross-section, detail the experimental process to image the cross-section and show how computer generated holography can be used in this process. The procedure used to construct the computer generated hologram (cgh) is discussed.

3.2 THEORY

Consider an opaque three-dimensional body. We would like to obtain an idea of the density variation over a particular cross-section in the interior of the body. It is not possible to measure the absolute density variation of an interior section without cutting up the body. However, we can measure the variation in X-ray absorptivity over the cross-section

non-invasively. The absorptivity of X-rays and the absolute density are quite closely related, so that knowledge of the X-ray absorptivity effectively gives us the required information.

Let the variation in X-ray absorptivity across a thin cross-section be denoted by $\omega(x,y)$. If a wavefront of X-rays with an intensity i_o is incident on the edge of the cross-section from the y direction, that is the X-rays pass through the section, then the energy transmitted through the section is E(x) given by

$$E(x) = i_o \, e^{-\int_{-\infty}^{\infty} \omega(x,y)\,dy} \tag{3.1}$$

The X-rays that pass completely through the section are absorbed by a film plate placed behind the body parallel to the x-axis. The film is developed so that the emulsion records the logarithm of the energy that strikes the film. If the photographic density along the plate is f(x) then

$$f(x) \propto \int_{-\infty}^{\infty} \omega(x,y)\,dy \tag{3.2}$$

The developed photographic plate is often called a shadowgram or projection. If the direction of the incident X-rays is not parallel to the x axis, then eqn (3.2) is redefined to allow for an arbitrary orientation between the cross-section and the X-ray beam.

Take a new system of cartesian coordinates r,s with the same origin as for the x,y coordinate system. The angle between the x axis and the r axis is θ. The direction of the

incident X-ray beam is along the s axis and the film plate is behind the body parallel to the r axis. All these geometrical relations are shown in Fig. 3.1.

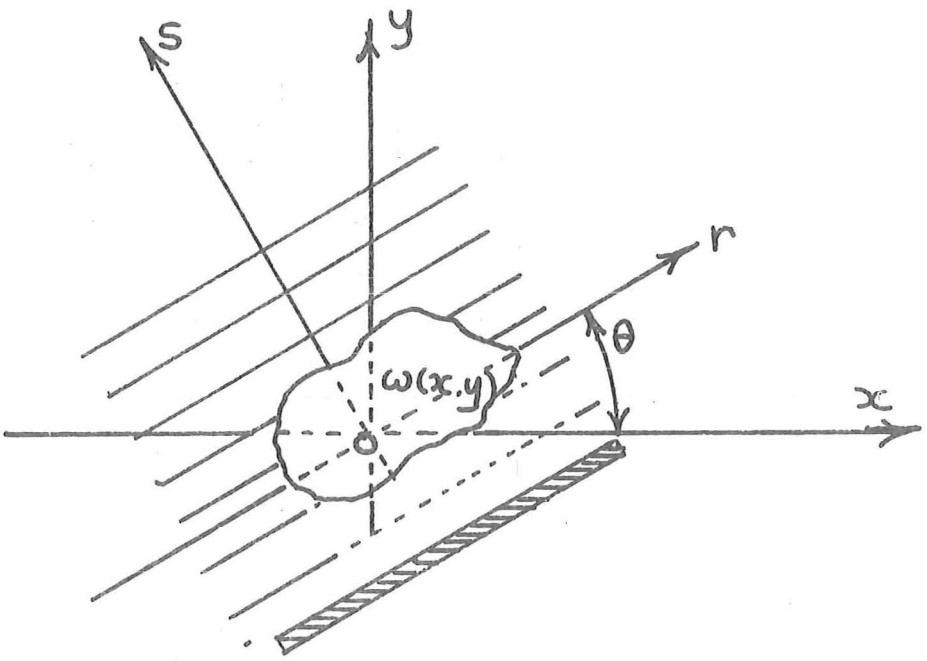


Fig. 3.1 Coordinate system of the object space.

Thus the developed density of the photographic plate due to incident radiation from the s direction is $f(r,\theta)$ where

$$f(r,\theta) \propto \int_{-\infty}^{\infty} \omega(x,y)\,ds \tag{3.3}$$

We define $\Lambda(\rho,\theta)$ as the one-dimensional Fourier transform of $f(r,\theta)$ with respect to ρ and r:

$$\Lambda(\rho,\theta) \longleftrightarrow f(r,\theta)$$

$$\therefore \quad \Lambda(\rho,\theta) = \int_{-\infty}^{\infty} f(r,\theta)\, e^{j2\pi\rho r} dr$$

$$= \iint_{-\infty}^{\infty} \omega(x,y)\, ds\, e^{j2\pi\rho r} dr$$

$$= \iint_{-\infty}^{\infty}\iint_{-\infty}^{\infty} \Omega(\alpha,\beta) e^{-j2\pi(\alpha x+\beta y)} e^{j2\pi\rho r} ds dr dx dy \quad (3.4)$$

The simplification of this expression is obtained by substituting in values for the various coordinates. Smith et al. (1973) develop this thoroughly, showing that

$$\Lambda(\rho,\theta) = \int_{-\infty}^{\infty} \Omega(\alpha,\beta)\, \delta(\eta)\, d\eta \quad (3.5)$$

where η is a coordinate of a cartesian coordinate system (ξ,η) that is related to the (α,β) coordinate system by rotation through the angle ϕ, as is shown in Fig. 3.2.

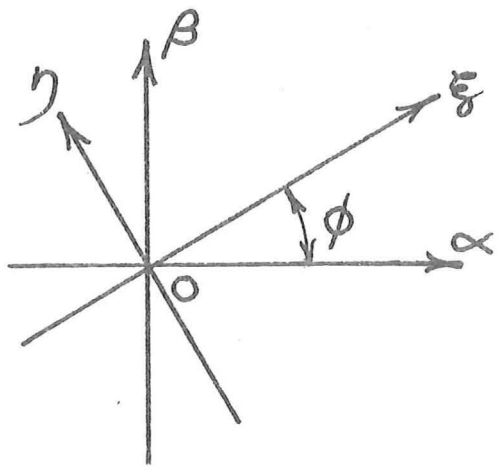


Fig. 3.2 Coordinate system in Fourier Transform space

From eqn (3.5), $\Lambda(\rho,\theta)$ is the value of the FT distribution $\Omega(\alpha,\beta)$ along a line running through the origin of coordinates at an angle θ to the α axis.

The cross-section in the object plane is now rotated a further angle $\Delta\theta$, so that the angle between the r axis and the x axis is $\theta+\Delta\theta$. The shadowgram of the cross-section in this orientation is measured to produce $f(r,\theta+\Delta\theta)$. Its one-dimensional FT is $\Lambda(\rho,\theta+\Delta\theta)$ which is the value of the $\Omega(\alpha,\beta)$ distribution along a line inclined at $\theta+\Delta\theta$ to the α axis. Thus, after two shadowgrams have been measured and their FT calculated we know the values of $\Omega(\alpha,\beta)$ along two radial lines. The cross-section is rotated in further steps of $\Delta\theta$ and the procedure repeated until the whole of $\Omega(\alpha,\beta)$ distribution has been sampled along a series of radial lines, all intersecting at the origin and with an angle of $\Delta\theta$ between each line. An interpolation procedure is then used (cf. Smith et al., 1973) to estimate the values between the radial lines. Once the whole of the $\Omega(\alpha,\beta)$ distribution has been estimated, the image of the cross-section is calculated using the two-dimensional FT.

It is now convenient to introduce some definitions. If $\omega(x,y)$ is the variation of X-ray absorptivity over the cross-section, then $\omega(x,y)$ represents the perfect image obtained if there are no imperfections introduced during the process. However, this cannot be obtained in practice. We call $\omega_N(x,y)$ the image obtained from N shadowgrams. Consequently, the FT of $\omega_N(x,y)$ computed by the procedure outlined above is $\Omega_N(\alpha,\beta)$.

It is instructive to consider the effect on the image $\omega_N(x,y)$ of the imperfect estimate $\Omega_N(\alpha,\beta)$ of the continuous distribution $\Omega(\alpha,\beta)$. Recall from Chapter 1 that an imaging system can be characterised by its psf and that the image can be written as a convolution of the psf and the best available image, which in this experiment is $\omega(x,y)$. Let psf_N be the psf of the process using N shadowgrams, then

$$\omega_N(x,y) = \omega(x,y) \otimes psf_N \tag{3.6}$$

Often it is convenient to use polar coordinates in which case we write $\omega(x,y)$ as $\lambda(r,\theta)$, so that eqn (3.6) becomes

$$\lambda_N(r,\theta) = \lambda(r,\theta) \otimes psf_N \tag{3.7}$$

Note also that if $\Lambda(\rho,\phi)$ is the FT of $\lambda(r,\theta)$, then

$$\Omega_N(\alpha,\beta) \equiv \Lambda_N(\rho,\phi) \tag{3.8}$$

Let us now calculate the psf_N. Consider a point source radiating in the object plane. A superscript † denotes all values pertaining to this object. Its continuous FT distribution $\Omega^\dagger(\alpha,\beta)$, or $\Lambda^\dagger(\rho,\phi)$, has a constant amplitude over the whole of the α,β, or ρ,ϕ, plane, and we may write $\Lambda^\dagger(\rho,\phi)$ in terms of a Fourier expansion

$$\Lambda^\dagger(\rho,\phi) = \sum_{m=-\infty}^{\infty} e^{jm\phi} \tag{3.7}$$

Now, $\Lambda_N^\dagger(\rho,\phi)$ consists of a collection of N radial lines equally spaced in angle having unit amplitude on the radial lines and zero amplitude between the lines. We may write $\Lambda_N^\dagger(\rho,\phi)$ in terms of its Fourier expansion

$$\Lambda_N^\dagger(\rho,\phi) = \sum_{m=-N}^{N} e^{jm\pi/N} \tag{3.8}$$

From Papoulis (1968) chapter 5, section 4 the two dimensional FT of $\Lambda_N^\dagger(\rho,\phi)$ is

$$\lambda_N^\dagger(r,\theta) = \sum_{m=-N}^{N} \Xi_m(r)\, e^{jm\theta} \tag{3.9}$$

where

$$\Xi_m(r) = \int_0^\infty J_m(2\pi\rho r)\rho d\rho \tag{3.10}$$

Let $\Lambda^\dagger(\rho,\phi)$ exist only for $\rho < A$. That is, the FT distribution is only calculated out to a radius A. Now,

$$\Xi_m(r) = \int_0^A J_m(2\pi r\rho)\rho d\rho$$

$$= \frac{J_1(2\pi Ar)}{2\pi Ar} \tag{3.11}$$

and eqn (3.9) may be written as,

$$\lambda_N^\dagger(r,\theta) \doteq \frac{J_1(2\pi Ar)}{2\pi Ar} \sum_{m=-N}^{N} e^{jm\theta}$$

$$= \frac{J_1(2\pi Ar)}{2\pi Ar} \frac{\sin(N+\frac{1}{2})\theta}{\sin(\theta/2)} \tag{3.12}$$

Since $\lambda_N^\dagger(r,\theta)$ is the distribution of light in the image plane due to a point source in the object plane then

$$psf_N = \lambda_N^{\dagger}(r,\theta) \tag{3.13}$$

Consequently eqn (3.7) may be written as

$$\begin{aligned} \lambda_N(r,\theta) &= \lambda(r,\theta) \otimes psf_N \\ &= \lambda(r,\theta) \otimes \lambda_N^{\dagger}(r,\theta) \\ &= \lambda(r,\theta) \otimes \left[\frac{J_1(2\pi Ar)}{2\pi Ar} \frac{\sin(N+\frac{1}{2})\theta}{\sin(\theta/2)} \right] \end{aligned} \tag{3.14}$$

which specifies the number of shadowgrams required to reconstruct the image to a required resolution.

3.3 EXPERIMENTAL PROCEDURE

We now apply the theory, outlined in section 3.2, to determine a cross-section in a real body.

Firstly, the body under investigation is placed so that its largest dimension lies along the z axis. For example, a cylindrical or nearly cylindrical object would have its axis of rotation along the z axis. Then, the particular slice or cross-section that is to be imaged is marked in some suitable fashion. (As a series of X-ray photographs is to be taken, it is essential to be able to measure the data from the same section in every photograph.)

The body is placed in an X-ray field and a shadowgram is made. The portion of the shadowgram that corresponds to the marked section is scanned and the density measured. This is $f(r,0)$. Then, the body is rotated a small angular increment

around the z axis, and a second shadowgram is taken and subsequently scanned. This is $f(r,\Delta\theta)$. This procedure is repeated until the body has been rotated a total of π radians. We now have a series of shadowgram scans all taken at slightly different angles.

Each shadowgram scan is digitised and fed into a computer. A linear Fourier transform is performed on each scan as it is read in and then this data is stored. Each of these Fourier transformed scans represents a radial line of known data in the unknown distribution $\Omega(\alpha,\beta)$. As each shadowgram scan is read in and Fourier transformed it is added to the data already in the α,β plane from the previous Fourier transformed scans. After all the shadowgram scans have been Fourier transformed and added into the distribution in the α,β plane, the required $\Omega(\alpha,\beta)$ distribution is approximated by a series of radial lines of known data.

Various methods (Smith et al., 1973) were tried to interpolate from these radial lines of data to get the full $\Omega(\alpha,\beta)$ distribution. However, there is often no visible difference in the final image between the simplest and most sophisticated interpolation procedures. Consequently, for all of the work reported here, a simple linear interpolation is used.

Once the full $\Omega(\alpha,\beta)$ distribution is determined, all that remains is to perform a two dimensional Fourier transform to obtain $\omega(x,y)$.

If this process is to become standard clinical, radiographic practice, it may be necessary to consider methods that do not involve the use of large memory digital computers.

This is because an average hospital may only have access to a medium sized computer.

Digital computation of two-dimensional Fourier transforms often takes an inconveniently long time, even using the Fast Fourier Transform algorithm. However, two dimensional transforms can be computed instantaneously by a coherent optical system, if the data is available on a transmissive recording medium.

To perform an optical two-dimensional Fourier transform on $\Omega(\alpha,\beta)$, the numbers stored in the computer that represent $\Omega(\alpha,\beta)$ must be converted into a variable density in a sheet of film. (The developed film then becomes the input to the optical Fourier transform.) As $\Omega(\alpha,\beta)$ is a complex distribution, it is necessary to calculate a hologram so that the data can be printed out and photographed.

A section of $\Omega(\alpha,\beta)$ is read into the core memory of the computer and $R(\alpha,\beta)$ added to it. The result of this addition is squared to give $I(\alpha,\beta)$:

$$I(\alpha,\beta) = |R(\alpha,\beta) + \Omega(\alpha,\beta)|^2 \qquad (3.15)$$

The hologram $I(\alpha,\beta)$ is then printed, or plotted out, depending on the type of cgh that is used. Eventually the whole of $\Omega(\alpha,\beta)$ is converted into a hologram distribution, and printed out. This is photoreduced and the developed transparency used as a Fourier hologram. One of its reconstructed images gives $\omega(x,y)$.

3.3.1 The Computer Generated Hologram

For much of the work reported here it was convenient to use the IBM 360/44 at the University of Canterbury Computer Centre. The hologram was produced by the optical analogy method and the computer artwork was produced by the computer's line printer chosen mainly because of its printing speed.

The distribution in the hologram plane has an intensity $I(\alpha,\beta)$ given by eqn (3.15) and, after it has been scaled, the intensity lies between 0.0 and 1.0. These computed intensity levels are now printed out by the line printer. When a teletype or line printer character (A-Z, 0-9) is printed, there is a pattern of black carbon on the white computer paper. Consequently when viewed from a distance so great that the individual characters cannot be identified, there is a certain 'blackness' associated with each character depending on how much carbon the character prints on the paper. However, there is very little difference in 'blackness' between the character that covers the most area, i.e. H, and the character that covers the least area, i.e. -. Consequently, the intensity level at any sample point is simulated by overprinting several line printer characters. To do this, the character feed of the line printer is suppressed and the computer prints a sequence, or combination or characters in the same position until a desired level of 'blackness' is reached. Once this level is reached, the character-feed is incremented by one unit and the computer then goes through the same procedure for the next sample point. Once a line of the hologram is printed out the line-feed is incremented by one unit and the character-feed is re-set to zero. It was found that the maximum number of useful overprinted characters was eight, no matter which

characters were used. That is, no measurable increase in 'blackness' accrued by overprinting a ninth character.

After the complete hologram distribution is printed out on white paper, each sample point is represented by a partially inked in area at a fixed position. The perimeter of this area is determined by the outside edges of all the characters used at that sample point. When a group of sampled intensities is printed out and photographed, each partially inked in area becomes a small transparent aperture with an opaque background. For example, if the letter H is printed out, its subsequent photograph, when illuminated in an optical bench, has a transparent aperture which passes approximately 5/7 as much light as a photograph of the characters H, L, F all overprinted. The latter combination of characters looks like: ▣. Thus, the more that the area is inked-in by the line printer, the 'blacker' the combination appears, and a photograph of the combination passes more light, that is, it is more transmissive.

The difference in transmissivity between different combinations of characters is measured experimentally. The 'blackness' is determined by how completely the area covered by the characters is inked in, and as well, there are some more or less random effects, usually due to incorrectly set characters in the line printer. Thus, the transmissivity of each combination is measured by printing out, and then photographing a whole block of identical combinations. The total amount of light passed by any block is then divided by the number of combinations in that block. In this way, the average transmissivity of a particular combination is obtained.

The amount of light passed by a photograph of a block of similar combinations is measured by comparing it to the amount of light passed by a calibrated density step wedge. The calibrated density step wedge used was a Kodak No. 2. Figure 3.3 shows some of the character combinations used and they are arranged in ascending order of 'blackness'.

Once the transmissivities of various character combinations have been measured, it is necessary to select from them all a set that corresponds to the intensity levels between 0.0 and 1.0. It was found that no matter what combinations of character were used only nine sensibly different transmissivities could be found. When the transmissivities of these combinations are arranged in ascending order, as in Table 1, it is found that there is a slight tendency for the intervals between the transmissivity levels to be smaller around a transmissivity of 0.6.

Certain errors occur when any continuous distribution is quantised into permitted levels. The greater the interval between the levels, the greater is the error that is introduced. Thus, for a continuous distribution that extends between 0.0 and 1.0, if we are to quantise this distribution into permitted transmissivity levels (Table 1), then a greater proportion of errors will be introduced when the intensity of any sample tends to the values 0.0 and 1.0, since these are where the intervals between the permitted levels are greatest.

However, it can be seen from Fig. 3.4 that for a typical cross-section $\omega(x,y)$, the Fourier hologram produced by the procedures outlined tends to 0.0 or 1.0 only over a small proportion of its total extent. Thus, to minimise the errors

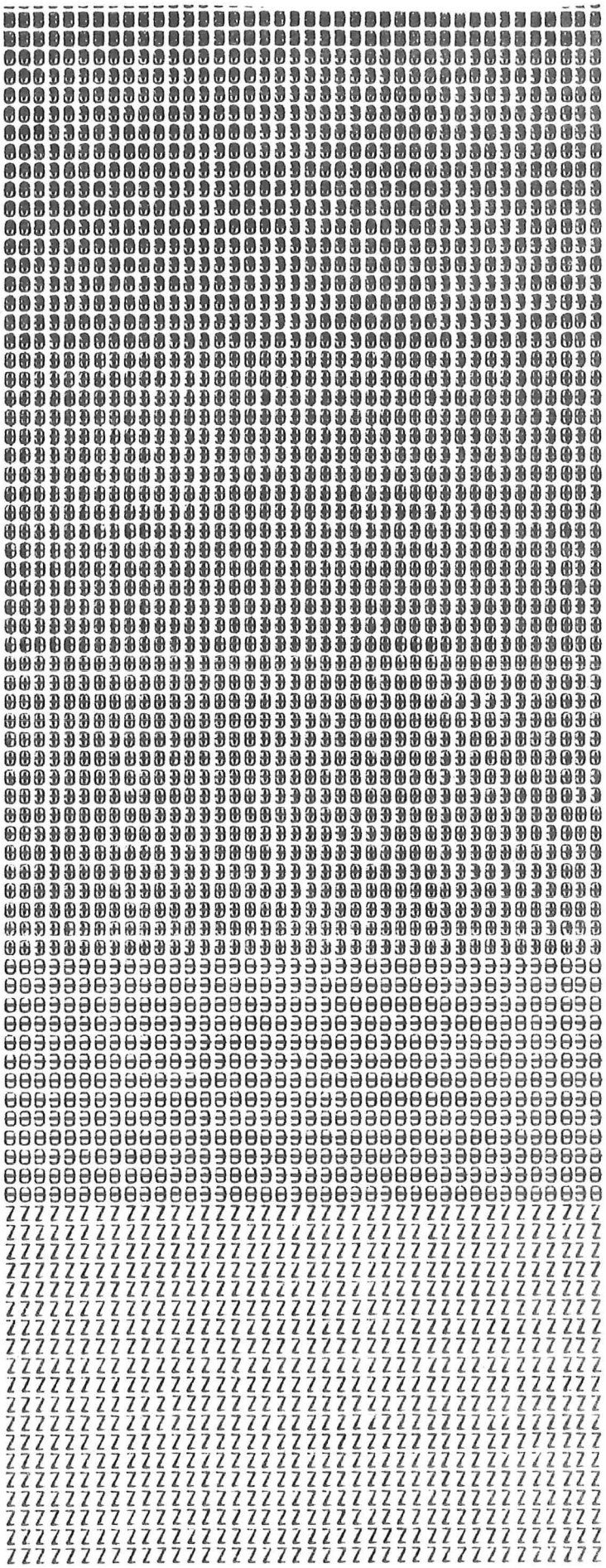


Fig. 3.3 Various character combinations used arranged in ascending order of blackness.

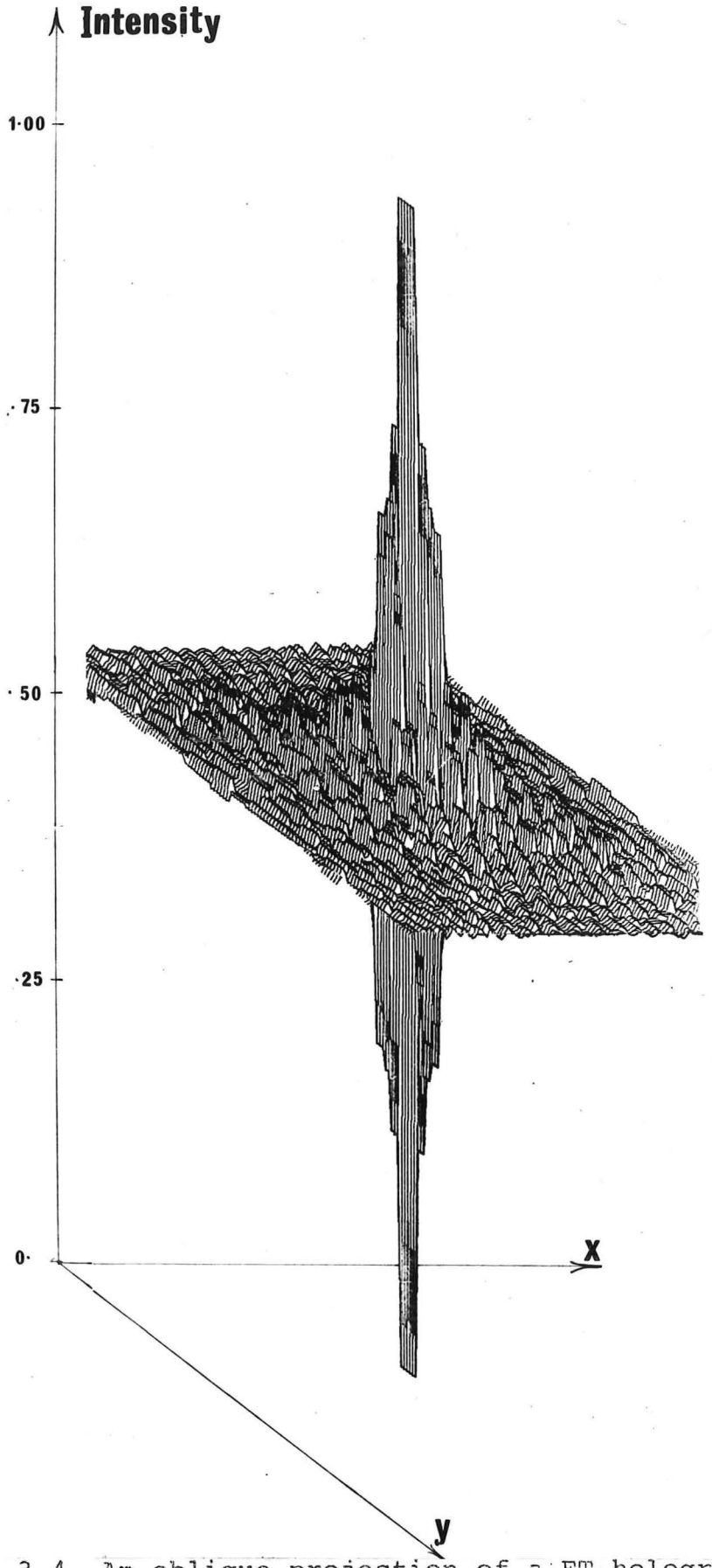


Fig. 3.4 An oblique projection of a FT hologram formed from a specular object.

TABLE 1

Transmissivities of overlaid teletype characters photographed on Kodak 649F film, processed in D-19 Developer.

Teletype characters	Transmissivities, when the teletype characters are printed on standard white, computer print-out paper.
No characters	0.33
z	0.45
O -	0.52
O +	0.55
O + ·	0.60
O + ·X	0.66
O + ·XB	0.71
O + ·XB'	0.74
O + ·XB'VA	0.84

introduced by having only certain transmissivities permitted, it is better to have smaller intervals between the permitted levels around 0.5 to 0.6. Taub and Schilling (1971) apply this type of reasoning to time signals. They use the term 'compounding' to describe permitted levels that have nonequal intervals between adjoining levels. Consequently, it would seem that less noise is introduced by quantisation when the interval between levels is smaller around a transmissivity of 0.6 than if all the levels were equally spaced. Although Table 1 is not an optimum from this aspect it represents the best selection of combinations of characters that we could obtain.

A limit on the usefulness of a hologram formed by the line printer is that the maximum number of samples in any one line is 128. This is the maximum number of characters in one line that the IBM 360/44 can produce. Thus, the largest size hologram that can be printed on a single sheet of paper has only 128 x 128 sample points.

It is possible to increase the extent of the hologram plane by printing out only one section of a larger hologram at one time, with each section having 128 x 128 samples. Once the whole of the larger hologram has been printed out, the separate pieces of paper that each section is printed on can be assembled together before they are photoreduced onto film. However, this solution is tedious and requires very careful aligning. A small misorientation of one section can give serious errors in the reconstructed image $\omega(x,y)$.

As an alternative to using the line printer to print out a larger hologram, a photographic facsimile machine was used. This has a larger number of resolvable sample points, and a

hologram with 512 x 512 sample points could be made at one time, each sample point having 20 permitted transmissivity levels.

However, it is not possible at Canterbury to connect the IBM 360/44 computer directly to our Muirhead facsimile machine. Instead, the data is routed from the IBM 360/44 to an EAI 590 hybrid computer in the Electrical Engineering Department. The hybrid computer then drives the facsimile machine via its analogue to digital converters. The programs to interface the two computers and to drive the facsimile machine were written by W.K. Kennedy and Dr T.M. Peters.

The technique used to make the facsimile machine print out an optical analogue hologram is essentially the same as the technique using the line printer. However, instead of a combination of overprinted characters to give the intensity at a sample point, a modulated light in the facimile machine is used.

The facsimile machine is set to print the hologram plane out over a 7" by 5" area. As the light pen rapidly moves over a sheet of photographic paper, it glows with more or less intensity depending on the intensity of the hologram plane that corresponds to the position of the pen. When the pen has printed out the hologram distribution, the photographic sheet is removed from the facsimile machine and is developed. In all the experiments reported here the photographic paper used was Agfa BS1 developed for two minutes with intermittent agitation in Bromophene at 68^{o}C. This hologram representation is photoreduced onto a transparency as before.

As with the lineprinted hologram, the transmissivity of any level produced by the light pen was determined experimentally. A selection of levels is shown in Fig. 3.5. Various different progressions from white to black are shown. The progression that looks most linear (or even most pleasing) to the eye, is not necessarily the most linear one. The transmissivities that were used to produce the results are shown in Table 2.

Various papers have been written on grey-scale simulation by computer output peripherals (cf. Knowlton and Harmon, 1972, Perry and Mendelson, 1964) and all these schemes may be adapted to produce cgh's of the optical-analogue type.

3.4 RESULTS

Two separate bodies were used to test the experimental procedure. The first body is a hollow aluminium cylinder. As any cross-section is circularly symmetric, only one projection or shadowgram is required, since all the other shadowgrams are identical. The exterior and interior diameters of the hollow cylinder were used as input data and the shadowgrams calculated using a program provided by Dr T.M. Peters. Since the $\Omega(\alpha,\beta)$ distribution is calculated completely digitally, $\Omega(\alpha,\beta)$ may be considered effectively exact, so that the quality of the image is an index of the accuracy of the cgh. The Fourier hologram of this body resembles Fig. 3.4 in a side looking view of the normalised intensity distribution. The hologram shows a large central peak and rapidly decreasing sidelobes.

The intensity level at each sample point in the hologram is then set to its nearest permitted level. There are nine permitted levels between 0.0 and 1.0 in a hologram produced by

TABLE 2

Transmissivities of the output levels using the facsimile machine. The photographic paper used in the machine was Agfa BS1 processed in Bromophene. The film used to photoreduce the artwork was Kodak 649F developed in D-19 to usual holographic principles (Appendix B).

Output level	Intensity in hologram plane	Transmissivity of developed cgh
1	0.05	0.25
2	0.10	0.30
3	0.15	0.35
4	0.20	0.39
5	0.25	0.43
6	0.30	0.47
7	0.35	0.51
8	0.40	0.55
9	0.45	0.58
10	0.50	0.60
11	0.55	0.61
12	0.60	0.63
13	0.65	0.66
14	0.70	0.69
15	0.75	0.72
16	0.80	0.76
17	0.85	0.80
18	0.90	0.84
19	0.95	0.88
20	1.0	0.92

the line printer. If the intensities in the hologram illustrated in Fig. 3.4 are divided into nine permitted levels, the nearest permitted level for most of the hologram is that with a transmissivity of 0.6. Consequently, much of the high frequency information in the hologram is lost since the variations in the high spatial frequencies are not great enough to jump from one permitted level to another.

To show the effect of the lack of high frequencies in the hologram, a hologram distribution was calculated of the hollow cylinder. A cross-section of this body is illustrated in Fig. 3.6(a). The hologram was inserted in the front focal plane of a Fourier transform lens and illuminated with coherent light. One of its reconstructed images is shown in Fig. 3.6(b). The indistinct edges that characterise this image are the result of having no high spatial frequencies in its hologram.

In an attempt to improve the definition of the edges of the reconstructed image, another hologram of the same cross-section was calculated. This time, however, the intensities were expanded in a non linear fashion. The maximum and minimum values remained as they were for the linear hologram, but the high spatial frequencies were considerably expanded. The intensities of the higher spatial frequencies now varied from 0.45 to 0.75, whereas in the linear hologram the intensities of the higher frequencies varied from 0.58 to 0.64.

The effect of this deliberate nonlinearity was to emphasize the edges of the reconstructed image to the detriment of the overall grey level. This shows up as a 'halo' around the edges of the reconstructed hollow cylinder, as shown in Fig. 3.6(c).

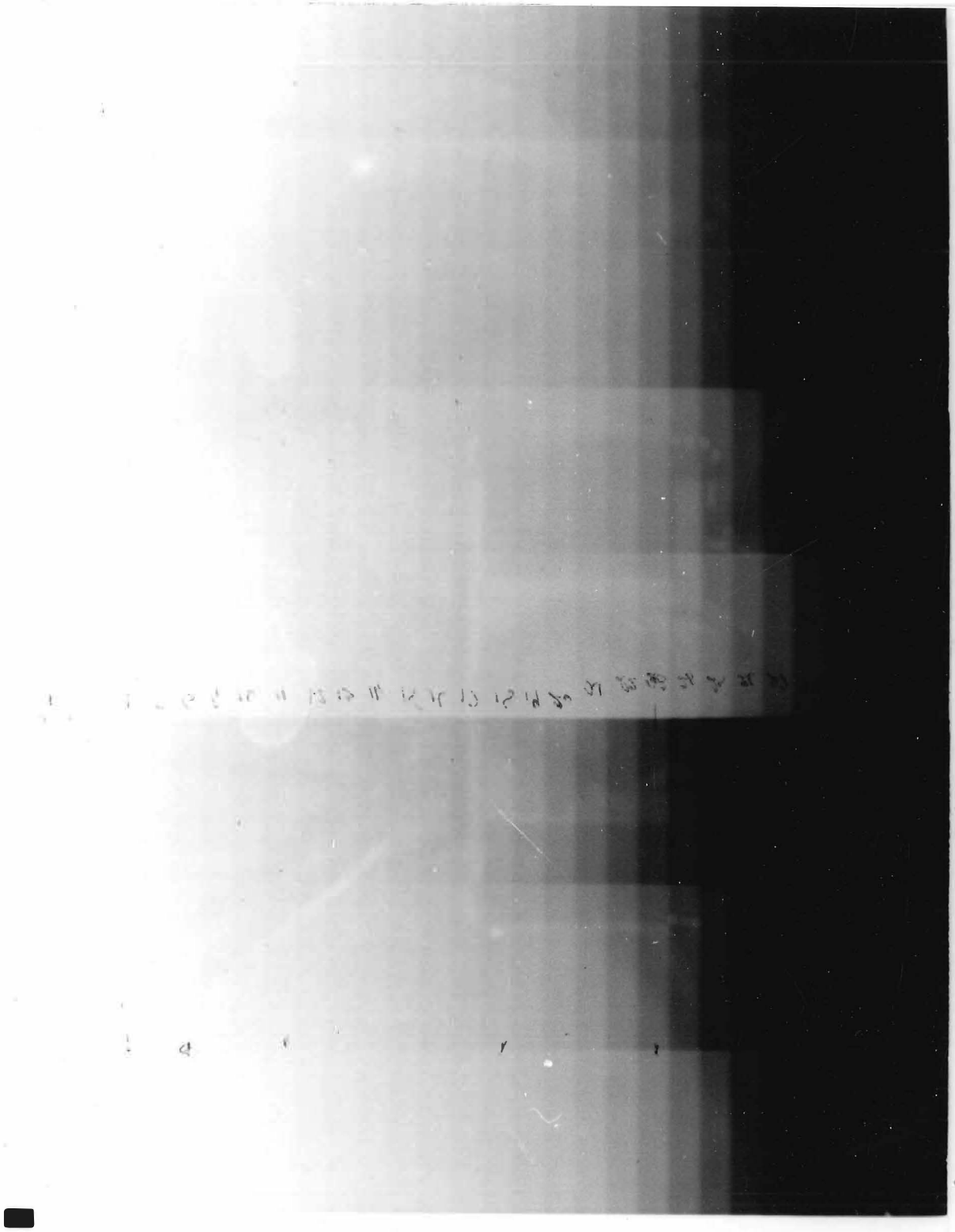
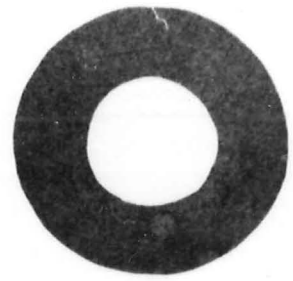
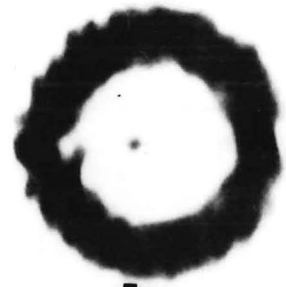


Fig. 3.5 Various intensities printed by the facsimile machine



a



b

c

Fig. 3.6 (a) The best available image of a circularly symmetric cross-section.
(b) A cross-section reconstructed from a line printer hologram
(c) A cross-section reconstructed from a line printer hologram with high frequency emphasis.

The second object used to test the procedure was a length of bone that had approximately the same dimensions as the hollow cylinder. The object was to obtain a clear image of a cross-section somewhere along the length of the bone using cgh's. The $\Omega(\alpha,\beta)$ distribution was calculated by Smith et al. (1973) from the shadowgrams made by Dr T.M. Peters. The procedure they used is briefly outlined. The position of the particular cross-section in the bone was marked, placed in an X-ray field, and a shadowgram was taken. The portion of the shadowgram that corresponded to the marked section was scanned, digitised, Fourier transformed and stored as outlined in section 3.3. Then the bone was rotated 9 ^{o} around its major axis, and a second shadowgram taken and processed. After twenty shadowgrams had been taken, the body had been rotated 180^{o}. After all the shadowgrams had been processed, the Fourier transform distribution $\Omega(\alpha,\beta)$ was approximated by the interpolation procedure given by Smith et al. (1973).

The intensity of a hologram distribution was calculated and printed out by the line printer, and by the facsimile machine. The two printouts were photoreduced onto Agfa Copex and developed in Rodinol 1/25 for nine minutes. Fig. 3.7(b) shows the right hand side of the line printer hologram.

The transparency of the line printer hologram was placed in the front focal plane of a Fourier transform lens and illuminated with spatially coherent light. One of its reconstructed images is printed out in Fig. 3.8(a). The transparency of the facsimile machine hologram was placed in the optical bench. One of its reconstructed images is printed in Fig. 3.8(b).

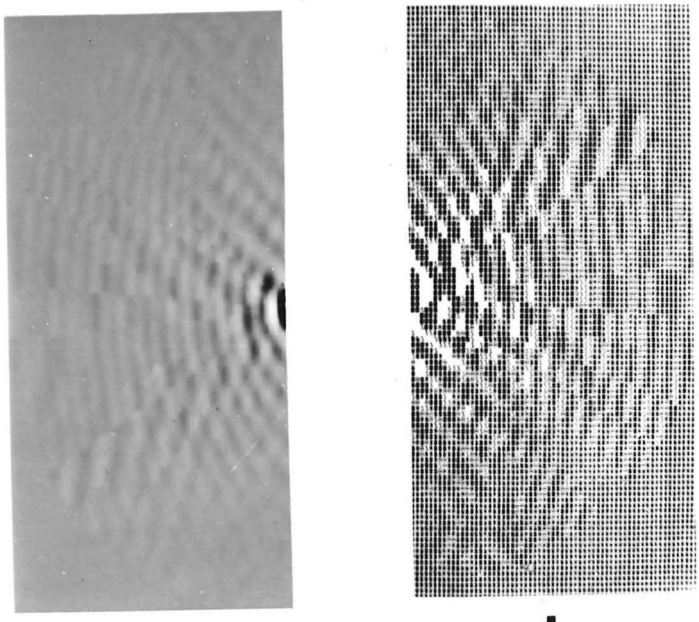


Fig. 3.7 Two computer generated holograms

(a) Left half plane of a hologram constructed using the facsimile machine.

(b) Right half plane of a hologram constructed using the line printer.

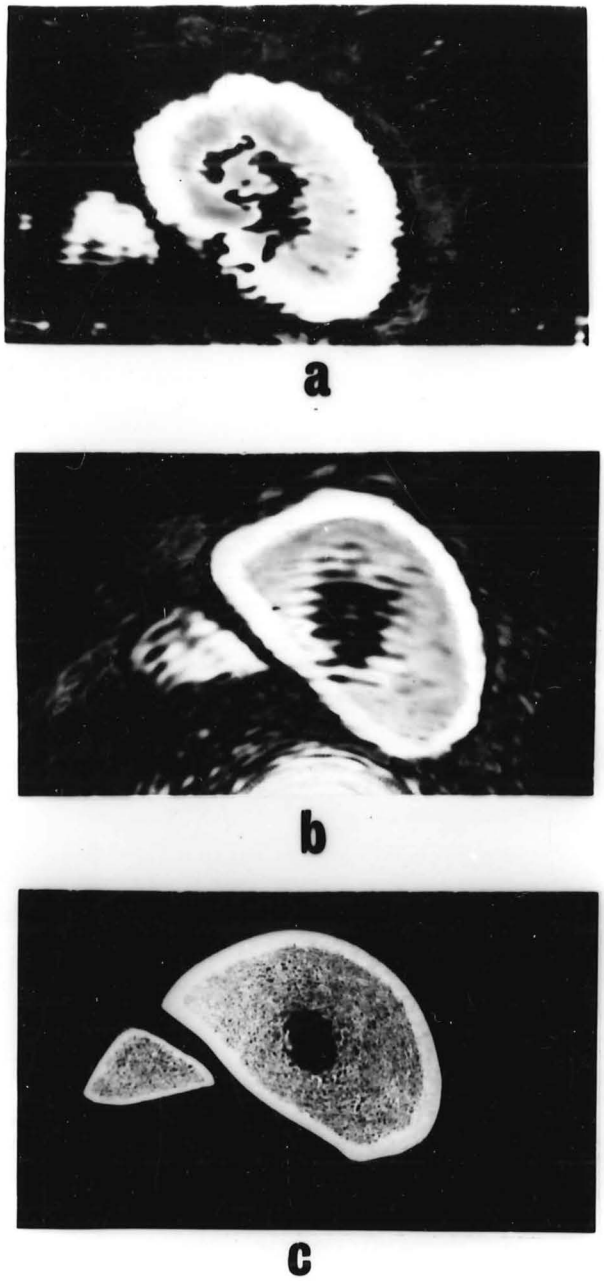


Fig. 3.8 (a) A cross-section of a bovine shin bone reconstructed using a line printer hologram.
(b) A cross-section reconstructed using a facsimile machine hologram.
(c) The cross-section cut out from the bone for comparison.

Subsequent to this experiment, the bone was removed and the marked cross-section sawn out. This is illustrated in Fig. 3.8(c). The comparison of Fig. 3.8(b) and Fig. 3.8(c) shows that the gross features of the marked cross-section have been faithfully recorded, but that there is a certain lack of detail. However, this procedure may still be a useful diagnostic aid.

3.5 DISCUSSION

The greatest problem in obtaining an accurate image is the limited dynamic range of both the printed output and the film. If a piece of pure white paper is photographed next to a piece of pure matt black paper the maximum difference between the transmissivity of the two parts in the correctly developed film is less than 20 db. This is because the black paper reflects somewhat more than 1% of the incident light, (assuming the white paper reflects 100% of the incident light). Todd and Zakia (1969) have shown that this is a major limitation in the fidelity of photographic prints, as a general outdoor scene has a range between the darkest of shadows to the brightest of highlights of up to 30 db.

Thus, if a hologram is computed and has important features that are separated in intensity by more than 20 db, it is not possible to reproduce all of them on a paper print.

The restriction in dynamic range is increased further when the printed hologram is transferred onto photographic film. When a film is used to record the fringe pattern of a hologram, its maximum dynamic range is 12 db. This restriction has been solved in part by Stroke (1969). He

places two unequally exposed films together. One film records the higher intensity levels of a specular Fourier hologram and the other film records the lower intensity fringes, which usually preserve the higher spatial frequencies.

Experimenters in computer generated holography have previously circumvented these restrictions in dynamic range by ensuring that their computed holograms do not have a large dynamic range in intensity. They take the wanted cross-section stored in the computer, and artificially impose a random phase across the object. This is directly analogous to placing a ground glass screen behind the object transparency in an optical Fourier holography experiment. However, in this type of solution to the dynamic range restriction, it is always necessary to know the cross-section or object distribution before the hologram is computed. Obviously, in the experimental procedure described in this chapter it is not possible to determine the cross-section before computing the hologram. If the cross-section were known, it would not be necessary to compute the hologram at all.

It may also be noted from Fig. 3.8(a) that the image produced from the line printer hologram is slightly inferior to the image from the facsimile machine hologram (Fig. 3.8(b)). A possible reason is that the facsimile hologram produces fewer multiple images due to the sample at any one point on the hologram being diffused and 'softened' slightly by the photographic procedure required by the facsimile machine output. Since there are fewer multiple images due to the facsimile hologram, the amount of light available for each image is slightly higher. For instance, in the image plane of

the line printer hologram there were approximately forty identical images. Thus any image selected may have less than 2% of the available light. In contrast to this, the image plane of the facsimile machine hologram usually has only two identical images. This means that the image selected has approximately 25% of the available light (the undeflected zero order has 50% of the available light). Thus the signal to noise ratio in the image from the facsimile machine hologram (Fig. 3.8(b)) is higher than in the image from the line printer hologram (Fig. 3.8(a)). Notice that although the edges of the image in Fig. 3.8(b) are clearer there is more interference from the undeflected beam from the laser. This shows as the concentric line centred about (b) just under the figure. This may be attributable to two causes. The first is that the lenses being used were not designed for monochromatic operation as Fourier transforming lenses. They were originally designed for photographic objectives, and so are coated and ground so as to minimise the worst aberrations of a multifrequency lens. Consequently, the lens design is not optimised for our use and the spherical aberration of the lens when used at full aperture is considerable. Tropel Inc. (U.S.A.) make lenses specifically designed for Fourier transform calculation, but these were not available to us.

Another possible cause of the unwanted interference is that the particular image is too close to the undeflected beam from the laser. Another hologram was calculated in which the reconstructed image was further away from the undeflected beam, but it was considerably poorer in definition than that shown in Fig. 3.8(b). However, the concentric lines were less

apparent in this image. Consequently, Fig. 3.8(b) represents the best image we were able to obtain of the cross-section being examined.

has been done on improvement of images from psv systems. One of the few examples is Robbins and Huang (1972).

Most research on image improvement has either assumed that the optical system is psi, or that the system can be forced by a series of mathematical operations to approximate a psi system. We limit this chapter to systems that may be considered psi.

Consider a distorted spatially incoherent imaging system that has an isoplanatic region in the object plane of greater extent than the radiating object. The system is termed psi and we write the image $i(x,y)$ as

$$i(x,y) = c(x,y) \otimes d(x,y) \tag{4.1}$$

where $c(x,y)$ is the best available image and $d(x,y)$ is the intensity psf. The FT of $i(x,y)$ is

$$I(\alpha,\beta) = C(\alpha,\beta)\, D(\alpha,\beta) \tag{4.2}$$

If the object is replaced by a point source radiator, $D(\alpha,\beta)$ can be determined, and so $C(\alpha,\beta)$ can be calculated from

$$C(\alpha,\beta) = I(\alpha,\beta)/D(\alpha,\beta) \tag{4.3}$$

The FT of $C(\alpha,\beta)$ gives the restored image $c(x,y)$.

A way of obtaining $c(x,y)$ from $i(x,y)$ is to use coherent spatial filtering. A transparency is constructed so that when it is illuminated by spatially coherent light, the amplitude distribution of light just behind the transparency is proportional to $i(x,y)$. This light passes through a lens and the distribution of light in the rear focal plane is proportional to $I(\alpha,\beta)$. A filter is introduced into this

focal plane so that the light that passes through it is multiplied by $1/D(\alpha,\beta)$. Consequently, the distribution of light just behind the filter is $C(\alpha,\beta)$. This light passes through a second lens to form an image $c(x,y)$ in the image plane. Fig. 1.2 shows a spatial filtering system.

The success of this method is determined by how accurately a filter can be made to produce $1/D(\alpha,\beta)$. It is convenient to define

$$Q(\alpha,\beta) = 1/D(\alpha,\beta) \tag{4.4}$$

In general $Q(\alpha,\beta)$ is complex so an optical filter that multiplies a wavefront by $Q(\alpha,\beta)$ is constructed holographically. The transmissivity of a hologram $t(\alpha,\beta)$ that multiplies an incident wavefront by $Q(\alpha,\beta)$ is

$$t(\alpha,\beta) = 0.5 + Q(\alpha,\beta) \text{ cosine } (\tfrac{\pi}{\ell})\alpha \tag{4.5}$$

where 'ℓ' is determined by the angle of the reference beam to the hologram plane. Since the transmissivity $t(\alpha,\beta)$ of any transparency is restricted to lie between 0 and 1, $Q(\alpha,\beta)$ must also be restricted in range, i.e.

$$|Q(\alpha,\beta)| < 0.5.$$

If $|Q(\alpha,\beta)|$ is greater than 0.5 then an inverse filter hologram cannot be constructed. For most distorted optical systems $D(\alpha,\beta)$ is zero for some (α,β) and so $Q(\alpha,\beta)$ is infinite at that point. Consequently holograms cannot be constructed to produce the ideal inverse filter $Q(\alpha,\beta)$. The best approximation to the ideal inverse filter $Q(\alpha,\beta)$ is

denoted by $\overline{Q(\alpha,\beta)}$ where

$$\begin{aligned} \overline{Q(\alpha,\beta)} &= Q(\alpha,\beta) && |Q(\alpha,\beta)| < 0.5 \\ &= -0.5 && Q(\alpha,\beta) < -0.5 \\ &= +0.5 && Q(\alpha,\beta) > 0.5 \end{aligned} \qquad (4.7)$$

There are three possible ways to produce an inverse filter hologram whose transmissivity is given by eqn (4.5) (where $Q(\alpha,\beta)$ is now replaced by $\overline{Q(\alpha,\beta)}$). These are the Vander Lugt filter, Stroke's two part filter, and a computer generated hologram.

The Vander Lugt filter is a FT hologram of a real object $\overline{q(x,y)}$ where

$$\overline{q(x,y)} \longleftrightarrow \overline{Q(\alpha,\beta)}$$

An essential restriction is that $\overline{q(x,y)}$ must be of limited extent. Many ideal filters $\overline{Q(\alpha,\beta)}$ have FTs $\overline{q(x,y)}$ that are complex and so it is not possible to construct a Vander Lugt filter.

Stroke's filter is made in two parts. We may write the function $\overline{Q(\alpha,\beta)}$ as having modulus and phase, i.e.

$$\overline{Q(\alpha,\beta)} = |\overline{Q(\alpha,\beta)}| \; e^{j \text{ phase } \overline{Q(\alpha,\beta)}}$$

Stroke first exposes a negative to $D(\alpha,\beta)$ and develops the negative to a gamma of 1 (see Appendix A). This results in a transparency that has a transmissivity proportional to $|\overline{Q(\alpha,\beta)}|$. Stroke then makes a hologram that records only the phase of $\overline{Q(\alpha,\beta)}$. The transparency and the hologram are then

stuck together to produce an approximation to the inverse filter hologram described by eqn (4.5). Stroke's filters are not easy to produce because the film available has limited dynamic range which means that $|\overline{Q(\alpha,\beta)}|$ is difficult to produce accurately. Stroke (1969) attempts to improve the dynamic range of the film by exposing two negatives to $D(\alpha,\beta)$. One negative records the high intensity peaks of $|D(\alpha,\beta)|$ and the other negative records the low intensity variations. Once these negatives are developed correctly, when stuck together they produce a transparency that approximates $|\overline{Q(\alpha,\beta)}|$. Another difficulty in constructing a Stroke filter is accurately locating $|\overline{Q(\alpha,\beta)}|$, on top of the hologram recording the phase of $\overline{Q(\alpha,\beta)}$.

The third way to produce an inverse filter holographically is using cghs. The computer is given the psf $d(x,y)$. It calculates $D(\alpha,\beta)$, then the inverse filter $\overline{Q(\alpha,\beta)}$, and finally the hologram $t(\alpha,\beta)$. Afterwards by a series of mechanical and photographic processes a transparency approximating $t(\alpha,\beta)$ is produced. The particular processes used depend on the computer peripherals available. Section 3.3 deals with the production of a particular type of cgh in detail.

4.3 IMAGES BLURRED BY CONSTANT LINEAR MOTION

There have been many photographs taken where the object scene has moved relative to the camera during the exposure of the film. Often these photographs are the result of an unrepeatable situation and so it is of interest to restore the unblurred scene from the blurred scene.

Consider a photograph of an object with a best available unblurred image of c(x,y). During the exposure of T secs there is relative, uniform movement between the camera and the scene of u ft/sec along the x axis. Harris (1966) shows that the image produced by the camera, which we call the blurred image i(x,y) is

$$i(x,y) = \int_{-T/2}^{T/2} c(x-ut,y)\,dt \tag{4.8}$$

For convenience let the extent of the motion during exposure be L, i.e.

$$L = uT$$

and writing eqn (4.8) in shorthand notation

$$i(x,y) = c(x,y) \otimes \text{rect}\ (\tfrac{x}{L}) \tag{4.9}$$

where

$$\begin{aligned} \text{rect}\ (\tfrac{x}{L}) &= 1 \qquad -\tfrac{x}{2L} < x < \tfrac{x}{2L} \\ &= 0 \qquad \text{Otherwise} \end{aligned} \tag{4.10}$$

As the orientation of the coordinates in the x,y plane is usually arbitrary, most examples of images blurred by constant linear motion can be forced into the form of eqn (4.9).

Since eqn (4.9) is in the form of a psi system we may use coherent spatial filtering to retrieve c(x,y) from i(x,y). The FT of eqn (4.9) is

$$I(\alpha,\beta) = C(\alpha,\beta)\ \frac{\sin(\pi L\alpha)}{\pi L\alpha} \tag{4.11}$$

Consequently the inverse spatial filter that restores the image is $\pi L\alpha/\sin(\pi L\alpha)$. However, if this filter is an optical hologram there are unavoidable restrictions as explained in section 4.2. If $\overline{Q(\alpha,\beta)}$ is the best approximation to the ideal filter function, it follows that

$$\overline{Q(\alpha,\beta)} = (\overline{\pi L\alpha})\ (\overline{1/\sin(\pi L\alpha)}) \tag{4.12}$$

where $|\overline{\pi L\alpha}| \leqslant .707$ (4.13)

and $|\overline{1/\sin(\pi L\alpha)}| \leqslant .707$ (4.14)

(so that $|\overline{Q(\alpha,\beta)}| \leqslant 0.5$).

Multiplying the FT distribution of the blurred image by the best estimate of the inverse filter gives

$$\begin{aligned}\overline{C(\alpha,\beta)} &= I(\alpha,\beta)\ \overline{Q(\alpha,\beta)} \\ &= I(\alpha,\beta)\ (\overline{\pi L\alpha})\ (\overline{1/\sin(\pi L\alpha)}) \\ &= C(\alpha,\beta)\ \frac{\sin(\pi L\alpha)}{\pi L\alpha}\ (\overline{\pi L\alpha})\ (\overline{1/\sin(\pi L\alpha)})\end{aligned} \tag{4.15}$$

It is convenient to define

$$C^{\dagger}(\alpha,\beta) = C(\alpha,\beta)\ \frac{\sin(\pi L\alpha)}{\pi L\alpha}\ \overline{\pi L\alpha} \tag{4.16}$$

so that

$$\overline{C(\alpha,\beta)} = C^{\dagger}(\alpha,\beta)\ (\overline{1/\sin(\pi L\alpha)}) \tag{4.17}$$

Note that if $C(\alpha,\beta)$ only exists for small α, eqn (4.16) reduces to

$$C^{\dagger}(\alpha,\beta) = C(\alpha,\beta) \sin(\pi L\alpha) \tag{4.18}$$

since

$$\overline{\pi L\alpha} = \pi L\alpha \tag{4.19}$$

for small α.

Consequently the FT of $C^{\dagger}(\alpha,\beta)$ is now

$$c^{\dagger}(x,y) = c(x,y) \times \frac{j}{2}\left[\delta(\tfrac{L}{2}) - \delta(\tfrac{L}{2})\right] \tag{4.20}$$

$$= \frac{j}{2} c(x-\tfrac{L}{2},y) - \frac{j}{2} c(x+\tfrac{L}{2},y) \tag{4.21}$$

Thus $c^{\dagger}(x,y)$ has two images, one centred at L/2 and one centred at -L/2. If the unblurred image is limited in extent so that the image at L/2 separates completely from the image at -L/2 then further processing to retrieve $\overleftarrow{C(\alpha,\beta)}$ from $C^{\dagger}(\alpha,\beta)$ required by eqn (4.17) is not necessary. The two images in $c^{\dagger}(x,y)$ separate only if the extent of the blurring L, is greater than the extent of the unblurred image. In the experiments reported in the following sections L is made greater than the extent of the unblurred image so that $c^{\dagger}(x,y)$ separates into two images. This makes the inverse filters slightly easier to construct but does not affect the generality of the procedures.

4.3.1 Optical Restoration of Images Blurred by Constant Linear Motion

An object of limited extent was photographed and during the exposure the object was moved along the x axis of the object plane a greater distance than the extent of the object (a picture of the stationary object is illustrated in Fig.

4.1(a)). The negative of the blurred image was reversed onto a second negative to form a positive transparency with an overall gamma product of 2. (Appendix A explains in more detail how film is processed to record images without distortion.) An enlargement of the |FT| of the blurred image is shown in Fig. 4.1(b). Note the vertical lines of zero intensity produced by the zeros in the transfer function $\sin(\pi L\alpha)/\pi L\alpha$. A cgh was computed to produce a transparency described by

$$t(\alpha,\beta) = 0.5 + 0.707\ (\overline{\pi L\alpha})\ \text{cosine}\ (\tfrac{\pi}{\ell}\ \alpha) \qquad (4.22)$$

where $(\overline{\pi L\alpha})$ has a restricted range expressed in eqn (4.13). A section of this cgh is shown in Fig. 4.2(a). The hologram transparency was placed in the filter plane of the spatial filtering system of Fig. 1.2, with a transparency of the blurred image in the object plane of the system. One of the two images in the image plane is shown in Fig. 4.2(b). The restored image reproduced in Fig. 4.2(b) is very poor and is little improvement on the blurred image, but it does represent the best restored image that could be produced from the blurred image using the cgh.

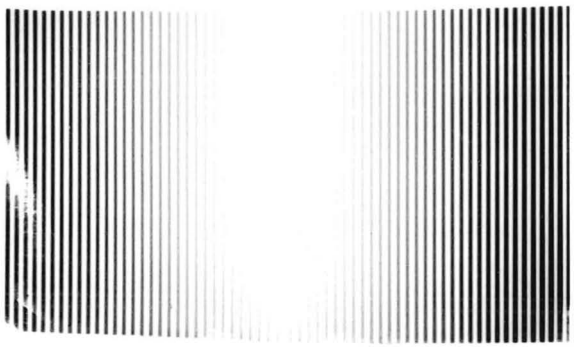
During the preceding experiment it was noticed that complete removal of all spatial frequencies for negative α produced some restoration of the blurred image. Although a filter which removes all frequencies for negative α may be produced by a cgh, it is easier to use a sharp edged opaque sheet such as a razor blade.

a

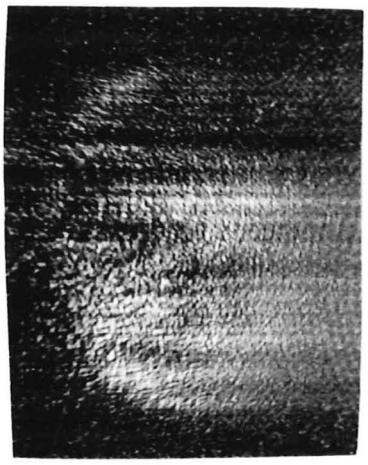
b

Fig. 4.1 (a) The unblurred object used to test the deconvolution procedures.

(b) A FT of the blurred image.



a



b

c

Fig. 4.2 (a) A section of a cgh used as a spatial filter.

(b) The filtered image using the cgh illustrated in (a).

(c) A reconstructed image resulting from a step function filter.

Firstly consider the effect of differentiation on the blurred image. The differential of eqn (4.9) is

$$i'(x,y) = c(x,y) \otimes \text{rect}' \left(\frac{x}{L}\right) \tag{4.23}$$

where the ' indicates differentiation

$$i'(x,y) = c(x,y) \otimes \tfrac{1}{2}\left[\delta\left(x - \frac{L}{2}\right) - \delta\left(x + \frac{L}{2}\right)\right]$$

$$= \tfrac{1}{2}c\left(x - \frac{L}{2}, y\right) - \tfrac{1}{2}c\left(x + \frac{L}{2}, y\right) \tag{4.24}$$

Note that the result of differentiating $i(x,y)$ has almost the same effect as mu iplying $I(\alpha,\beta)$ by α as shown in eqn (4.21). (There is a difference in phase but as it is usual to compare only the intensities of $i'(x,y)$ and $c^{\dagger}(x,y)$ the phase is immaterial.) Consequently, if some filtering operation results in the blurred image being differentiated, then there is some restoration of the unblurred image. However, as with eqn (4.21) there is the restriction that the amount of blurring L must be greater than the extent of the unblurred image for differentiation to produce two separated images.

Now consider the effect of totally removing all spatial frequencies for negative α on the blurred image $i(x,y)$. The spatial frequency distribution $I(\alpha,\beta)$ is multiplied by a unit step function to produce

$$Z(\alpha,\beta) = I(\alpha,\beta)\mathcal{H}(\alpha) \tag{4.25}$$

where $\mathcal{H}$ is unity for all positive α and zero otherwise.

Taking the FTs of eqn (4.25)

$$z(x,y) = i(x,y) \otimes \tfrac{1}{2}(\delta(x) + \frac{j}{2\pi x})$$

$$= \tfrac{1}{2}i(x,y) + \frac{j}{2\pi}\, i(x,y) \otimes \frac{1}{x} \qquad (4.26)$$

The success of this crude experiment depends on how closely convolution by 1/x in the second term of eqn (4.26) approximates differentiation. We assume that this approximation is valid and write

$$z(x,y) = \tfrac{1}{2}\, i(x,y) + \frac{j}{2\pi}\, i(x,y) \otimes \delta'(x)$$

$$= \tfrac{1}{2}\, i(x,y) + \frac{j}{2\pi}\, i'(x,y)$$

$$= \tfrac{1}{2}\, i(x,y) + \frac{j}{4\pi}\, c(x - \frac{L}{2},y) - \frac{j}{4\pi}\, c(x + \frac{L}{2},y) \qquad (4.27)$$

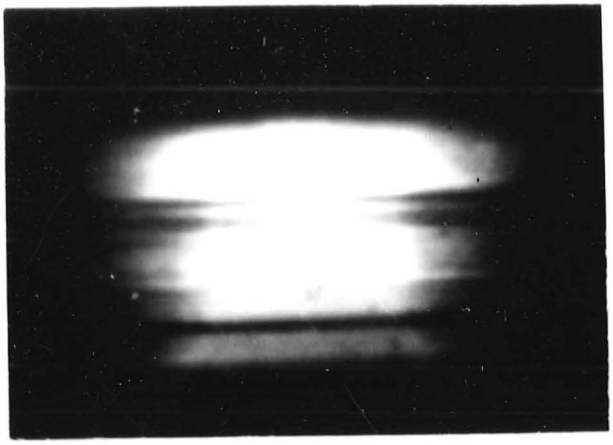
which suggest that in the image plane there is the blurred image $i(x,y)$, with the two unblurred images $c(x-\frac{L}{2},y)$ and $c(x+\frac{L}{2},y)$ superimposed over the top of the blurred image.

It is interesting to compare a restored image using a step function filter Fig. 4.2(c) and a restored image using a cgh filter Fig. 4.2(b). Since there is little visual difference between the two restored images and since the step function filter is far from mathematically ideal, the cgh is not very effective as an inverse filter. However, it may be that even if we could make a theoretically perfect cgh, that is, one with no restrictions on $t(\alpha,\beta)$, the inherent noise and non-linearities that occur when $i(x,y)$ is recorded would still prevent a worthwhile reconstruction.

4.3.2 Digital Reconstruction of Images Blurred by Constant Linear Motion

In an attempt to estimate the worth of the restoration scheme it was decided to simulate the optical restoration in a digital computer. As the transmissivity of a digital inverse filter is not limited to lie between 0 and 1 we can compute a theoretically perfect cgh. If the result of the digital restoration is no better than the result of the optical restoration, then the whole restoration procedure has doubtful value. If the result of the digital restoration is better than the optical restoration, then the cgh has doubtful value as an inverse filter.

The transparency that records the blurred image $i(x,y)$ was placed in the scanning plane of a computer controlled image scanner constructed by Kennedy and Peters (1974). After the blurred image was scanned, the computer stored a sampled quantised approximation to the blurred image. The FT of the sampled image was computed and multiplied to a theoretically perfect cgh. The cgh was calculated from eqn (4.22) and there were no restrictions in range placed on $t(\alpha,\beta)$ (or on $\pi L\alpha$). It was found by experimentation that if the transmissivity was not limited to some suitable maximum value, massive amplification occurred at large positive and negative α, which greatly increased the random grain noise in $i(x,y)$ and the restored image $c(x-\frac{L}{2},y)$ was buried in large random fluctuations. The most suitable maximum value of $|t(\alpha,\beta)|$ was found to be 200 although this varied considerably from picture to picture.



a

b

Fig. 4.3 (a) An image blurred by constant linear motion.
(b) A digitally reconstructed image.

The result of the digital deblurring scheme is illustrated in Fig. 4.3(b) where Fig. 4.3(a) is the blurred image. It can be seen that the digital restoration produces a superior image to the optical restoration. Consequently, it would seem that although the mathematics of the restoration scheme produce a reasonable image, the inherent restriction in transmissivity of an optical hologram make cghs unsuitable as inverse filters.

4.4 IMAGES BLURRED BY LINEAR MOTION THAT IS NOT CONSTANT

Consider a photograph of an object scene where there is non-constant relative motion between the camera and the scene during the exposure. In general, this motion is non-linear and variable. However, let us consider only linear variable motion along the x axis. If the motion may be described by u(t) then the film records a blurred image i(x,y) which may be written as a convolution of the unblurred image c(x,y), i.e.

$$i(x,y) = \int_{-T/2}^{T/2} c(x-u(t),y)\,dt \qquad (4.28)$$

If the rate of movement changes erratically during the exposure, then very little can be done to retrieve c(x,y) from i(x,y). However, if u(t) is monotonically increasing and more specifically if

$$u(t) = 1 + \sigma t + \frac{\sigma t^2}{2!} + \frac{\sigma t^3}{3!}$$

$$\doteq e^{\sigma t} \qquad (4.29)$$

then the FT of i(x,y) is

$$I(\alpha,\beta) = C(\alpha,\beta) \frac{\sin \pi(\alpha+j\sigma)L}{\pi(\alpha+j\sigma)L} \tag{4.30}$$

where L is the total extent of the motion during the exposure of T secs, i.e.

$$L = \int_{-T/2}^{T/2} u(t)\, dt \tag{4.31}$$

Alternatively, if u(t) is monotonically decreasing, eqns (4.29) and (4.30) still hold except that σ is now negative.

To retrieve $C(\alpha,\beta)$ from $I(\alpha,\beta)$ given by eqn (4.30) requires an inverse filter $Q(\alpha,\beta)$ where

$$Q(\alpha,\beta) = \pi(\alpha+j\sigma)L \; (1/\sin \pi(\alpha+j\sigma)L) \tag{4.32}$$

4.4.1 Optical Reconstructions

A cgh of the ideal inverse filter was constructed with the usual constraints applied to the allowable transmissivity. Unfortunately, the restored image was worse than the blurred image . This was not unexpected in light of the research performed in section 4.3.1.

From section 4.3 we can show that if L is greater than the extent of the unblurred image, a filter

$$Q(\alpha,\beta) = \pi(\alpha+j\sigma)L$$

should give some restoration. This filter was constructed using a cgh. Once again the restored image was worse than the blurred image. It is possible that these two filters produced some restoration, but the noise introduced by the cgh and the other necessary optical elements was far greater than any possible improvements.

We have also shown in section 4.3 that the differential of the blurred image often produces some restoration. Consequently, we now investigate a filter

$$Q(\alpha,\beta) = \alpha^2 \tag{4.33}$$

Note that this filter is purely positive real. Denote the result of the multiplication of this filter on $I(\alpha,\beta)$ by $\hat{C}(\alpha,\beta)$, thus

$$\begin{aligned}\hat{C}(\alpha,\beta) &= \alpha^2 \, I(\alpha,\beta) \\ &= (\alpha-j\sigma+j\sigma)^2 \, I(\alpha,\beta) \\ &= (\alpha+j\sigma-2j\sigma - \frac{\sigma^2}{\alpha+j\sigma}) \, C(\alpha,\beta) \, \frac{\sin \pi(\alpha+j\sigma)L}{\pi L} \\ &= (\alpha-j\sigma-\frac{\sigma^2}{\alpha+j\sigma}) \, C(\alpha,\beta) \, \frac{\sin \pi(\alpha+j\sigma)L}{\pi L}\end{aligned} \tag{4.34}$$

The FT of $\hat{C}(\alpha,\beta)$ given by eqn (4.34) is

$$\begin{aligned}\hat{c}(x,y) = \frac{j}{4\pi^2 L} \frac{\partial}{\partial x} \left[e^{\pi\sigma L} \, c(x-\tfrac{L}{2},y) - e^{-\pi\sigma L} c(x+\tfrac{L}{2},y) \right] \\ - \frac{\sigma}{\pi L} \left[e^{\pi\sigma L} c(x-\tfrac{L}{2},y) - e^{-\pi\sigma L} c(x+\tfrac{L}{2},y) \right] + \sigma^2 \, i(x,y)\end{aligned} \tag{4.35}$$

Since $i(x,y)$ is a smooth blurred image, it follows that $\hat{c}(x,y)$ is composed of the <u>differential</u> of the two <u>unblurred</u> images, centred at $+L/2$ and $-L/2$, superimposed on the two <u>unblurred</u> images and both of these are superimposed on the <u>blurred</u> image. Consequently a filter with a transmissivity proportional to α^2

gives some restoration of the unblurred image.

A cgh of this inverse filter was constructed and placed in the filter plane of a spatial filtering system with the blurred image as input. Although the restored image was not very clear, it was better than the restored images using the previous two optical filters. Since the filter expressed in eqn (4.33) was the most promising of all the optical filters used, it was decided to perform a digital simulation of a restoration scheme using this filter.

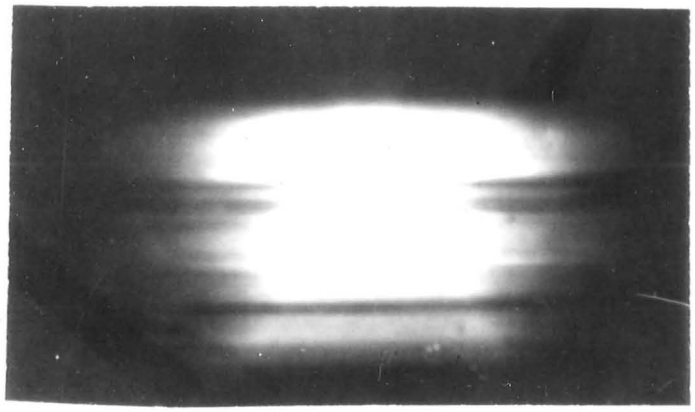
4.4.2 Digital Restoration

The photograph of the blurred image that was used as the input transparency to the optical restoration system was placed in the image scanner and a sampled quantised approximation to $i(x,y)$ stored in the computer. The FT of $i(x,y)$ was computed and this was multiplied to a digital filter $Q(\alpha,\beta)$ where

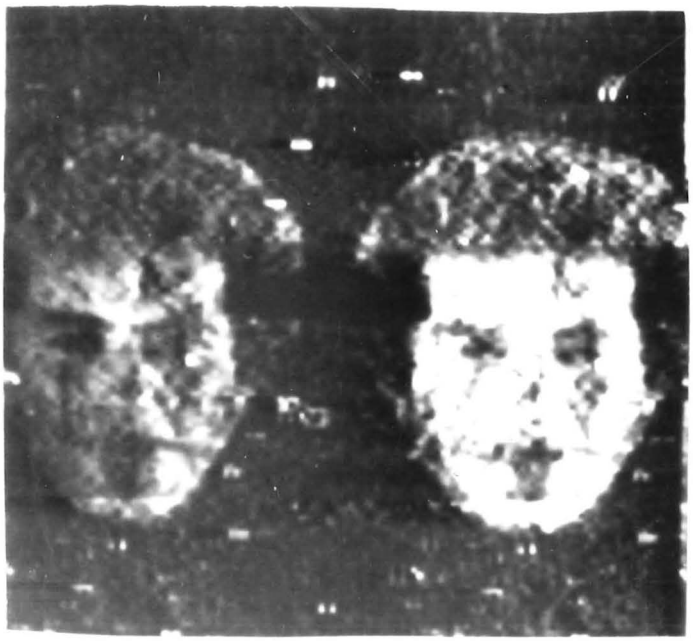
$$\begin{aligned} Q(\alpha,\beta) &= \alpha^2 && \alpha^2 < t_{max} \\ &= t_{max} && \alpha^2 \geqslant t_{max} \end{aligned} \qquad (4.36)$$

The absolute value of t_{max} is set by the programmer and varied considerably depending on the actual picture used as input.

Fig. 4.4(a) shows the image blurred by accelerated motion. The total extent of the blur L is set so that it is the same as the total extent of the blur in the example of uniform motion. Consequently, the image blurred by uniform motion (Fig. 4.3(a)) closely resembles the image blurred by non-uniform motion (Fig. 4.4(a)). The result of the digital restoration



a



b

Fig. 4.4 (a) An image blurred by accelerating linear motion.
(b) Two digitally reconstructed images.

of the image blurred by accelerated motion is shown in Fig. 4.4(b). The difference in intensities between the left side image and the right side image gives an estimate of the acceleration since the difference in intensities is given by $e^{\pi\sigma T}$ and $e^{-\pi\sigma T}$, as in eqn (4.31).

4.5 COMMENTS

Inverse spatial filtering inevitably becomes a comparison of digital and optical techniques. Since most inverse filters are complex, optical filtering usually requires holography. Holographic inverse filters are not easy to make optically since a well equipped optical laboratory is required. Since cghs require very little equipment not generally available, it would seem that cghs could become an easy way to produce inverse filters. From sections 4.3 and 4.4 it would seem that without sophisticated photographic techniques, cgh have limited applications as inverse filters owing to their limited aperture size and dynamic range. However, simple optical processing using cgh may well have application in preliminary estimation of parameters of degraded images, even though the actual restoration procedures are bound to be digital (Bates et al., 1974).

CHAPTER 5

SPECKLE INTERFEROMETRY AND HOLOGRAPHY

5.1 IMAGING THROUGH DISTORTING MEDIA

All forms of imaging systems distort the perfect image. The distortions may be introduced by a badly positioned lens, some phase distortion due to irregularities in the medium between object and image, or merely the finite size of the optical aperture.

Consider a coherent imaging system imaging a radiating object. This can be represented as a convolution of the best available image b(x,y) with the point spread function h(x,y) of the imaging system;

$$g(x,y) = b(x,y) \otimes h(x,y) \tag{5.1}$$

The purpose of any image restoration scheme is to attempt to remove the effect of the distortion h(x,y), and to find the resulting best available image b(x,y). In almost all cases of distorted imaging, the convolution due to the limited size of the lens has much less effect than the convolution due to other distortions.

Various methods have been advanced to remove the imperfections introduced by a distorted imaging system, and to obtain the best possible image. They fall into two major groups depending on whether the imaging system can be considered coherent or incoherent, and into several subgroups depending on whether the distortions are constant, periodic or random with respect to time.

5.1.1 Distorted Coherent Imaging

A coherent radiating object produces a set of radiating wavefronts which are distorted by the imaging system and these distorted wavefronts form a distorted image. Employing coherent radiation, it is possible to record the complex wavefront holographically. Later, this wavefront can be reconstructed to form the image. If the distorting medium does not vary with time and can be moved from one experimental set-up to another, it is often feasible to record the complex wavefront through the distorting medium and then reconstruct the wavefront through the same distorting medium. This way, in the progression from the object to the image, the wavefronts pass through two equal and opposite distortions, so that the final image is distortion free in principle. Both Leith and Upatnieks (1966) and Kologelnik (1965) use this technique for distortion free imaging.

Another holographic technique for distortion-free imaging has been advanced by Goodman et al. (1966). Both the reference wavefront and the object wavefront pass through the distorting medium, so that, if the object and the point source of the reference radiation are quite close together, the phase retardation of the object wavefront at any one point in the hologram plane is the same as the phase retardation of the reference wavefront. Thus, the position of the fringes in the hologram due to the combined wavefronts is not affected by the distorting medium. Because the position of the fringes does not shift, the phase distortions may vary during the exposure of the hologram. An essential restriction of this technique is that the distortion suffered by the reference wavefront

incident at a particular point in the hologram plane must be the same distortion that affects the object wavefront at that point. This means that both the reference source and the object must be within the isoplanatic region. As the volume of distortion becomes thicker the isoplanatic region in the object plane becomes smaller, and it is often difficult to fit both the source of the reference wavefronts and the object within this region. It is usually easier to restrict the distorting medium to a thin planar distribution just in front of the hologram, which gives the largest possible isoplanatic region.

Another holographic technique for distortion-free imaging may be used when the distorting medium does not vary in time. Firstly, a hologram is made recording the distorted wavefronts radiating from an unresolvable point source. Then this unresolvable source is replaced by the extended object and this object is viewed through the hologram. In this technique the hologram acts as a type of spatial filter correcting the distortions in the object wavefront. Upatnieks et al. (1966) uses this technique to correct some lens aberrations.

5.1.2 Distorted Incoherent Imaging

The correction of distorted images formed by incoherent illumination is the most common problem facing all users of optical systems. There are two possible approaches to improving these distorted images. The first is to compensate for, or try to eliminate, all possible forms of aberrations that affect the optical system. A common example of pre-detection compensation is correcting Seidal aberrations of a glass lens system using geometrical optics. There are many

papers and books on this subject (cf. Conrady 1929, 1960, Hopkins 1950, 1966, Wynne,1959).

A second approach to the problem of correcting a distorted image is to measure the affect of the distortions, so that the distorted image may be corrected after the image has been formed. This technique is usually called post-detection compensation and it does have several disadvantages. It is not often possible to form a 'real-time' corrected image and the corrected or restored image may take some time to produce. Another disadvantage is that some forms of distortion eliminate vital information. No amount of subtle mathematics can restore the lost information although it can suggest information that may have existed. This can produce misleading results on occasion, since the researcher may alter parameters to put in information that he would most like to see. This is a very real danger with iterative image correcting procedures. On each iteration the image looks more and more as the researcher desires. An example of this type of image enhancement is given by MacAdam (1970). However, despite these disadvantages, post-detection image restoration and enhancement procedures are the key to modern image processing.

There are many techniques in use which attempt to correct a distorted incoherent image after the image has been recorded, and it is the type of distortion present in the image that determines the best technique in any particular problem.

Consider an image formed by an optical system where the distortion does not vary with time. McGlamery (1967) first determines the psf of the system by imaging a small point in

the object plane. The intensity of the image of this point object is measured. This is the psf of the distorted system. Then the point object is replaced by a more general spatially incoherent object and the distorted image is photographed and is processed by the inverse of the psf of the distortion. However, this technique requires a computer with a large amount of random access core memory, to be efficient.

Horner (1970) improves a distorted image by using a least-mean-square-digital-filter and Russel and Goodman (1971) claim that this is a superior technique to the inverse psf filter approach. The latter two researchers use a combined pre and post-detection image enhancement scheme. They deliberately introduce a filter into the lens aperture that distorts the image in some predetermined way, and this makes post-detection enhancement easier.

Mueller and Reynolds (1967) assume that the image forming system is stationary and linear, so that optically dividing the degraded image spectrum by the degraded transfer function suppresses the degradation. This can only be done if the degraded transfer function does not fall below the noise level for any spatial frequencies of interest.

5.1.2(a) Speckle Interferometry

Labeyrie (1970) proposes a post-detection enhancement technique with specific application to astronomy. The resolution of a large aperture telescope is usually limited by the seeing conditions of the atmosphere, rather than the diffraction limit of the telescope. The seeing conditions are determined by the amount of turbulence in the earth's atmosphere. Poor seeing conditions occur when the minor variations in refractive index due to the turbulent atmosphere

change completely in a short time, perhaps 1/100 sec. The instantaneous image of a star in a telescope limited by seeing conditions is composed of randomly spaced speckles of intense light over a darker background (Texereau, 1963). The size of each individual speckle is approximately the same size as a diffraction limited image, that is, the smallest image that the finite aperture telescope is capable of imaging. The positions of the speckles in the image plane change at the same rate as the atmospheric turbulence changes and this is the reason for stars appearing to 'twinkle' when viewed on a clear night.

When a normal photograph of a star is taken, the film emulsion is exposed for a time much longer than the redistribution time of the atmosphere, which is defined as the time taken for the turbulence in the atmosphere to change completely. So that, the image recorded by a long-time-exposure photograph is the average intensity of many instantaneous speckle patterns. If the film is exposed for a considerable length of time then the image becomes a smooth circular distribution and is called the seeing disc. Information that is larger than the diffraction limit but smaller than the seeing disc is usually lost.

Labeyrie's technique records stellar information down to the diffraction limit and is independent of the seeing conditions. A disadvantage is that the technique can only be used to resolve stars, i.e. measure their angular diameters. Labeyrie's process produces only the autocorrelation of the star and information about the star that can be obtained from the autocorrelation includes the diameter, oblateness and limb darkening (Gezari, et al. 1972, Bonneau and Labeyrie, 1973).

Labeyrie's technique is to take a series of short-time-exposure-photographs of a small star through a band-pass filter. The short time exposure ensures that the speckles do not move during the exposure, and the filter ensures that the optical wavelengths are restricted to a narrow bandwidth which emphasises the speckles in the image. The first star that is photographed is smaller than the diffraction limit. Labeyrie calls each speckle image a "speckle interferogram". The power spectra of all the speckle interferograms are averaged, where the power spectrum is defined as the squared modulus of the FT of a speckle interferogram. The averaged power spectra of all the speckle interferograms we call a speckle transform.

A series of speckle interferograms is now taken of a second star which is larger than the diffraction limit but smaller than the seeing disc. The speckle transform from all these speckle interferograms is compared with the speckle transform derived from the speckle interferograms of the small star. The relative sizes of the two speckle transforms gives the diameter and oblateness of the larger star. The limb darkening is found by repeating the whole procedure for a band pass filter centred at a different mean frequency. The difference in measured diameters at different optical frequencies is the limb darkening.

Gezari et al. (1972) use Labeyrie's technique to determine the essential details of a few previously unresolved stars, as well as to confirm details of stars measured by more conventional techniques. Harvey and Breckinridge (1973) use the technique to measure the distribution of spatial frequencies of the sun's surface.

Dainty (1973) shows that in theory it is possible to obtain diffraction limited information from low quality telescopes that, owing to aberrations, are incapable of resolving even the seeing disc. This is achieved by placing an additional random diffuser across the optical aperture. An extension of Dainty's work is that it may be possible to synthesise a large optical aperture by a two dimensional array of smaller elements.

Korff (1973) analyses Labeyrie's technique by assuming that the turbulence obeys a specific mathematical model, and Korff et al. (1972) give an heuristic analysis of the statistical distribution of the speckles in the image and in the FT of the image.

Knox and Thompson (1973) tried to extend Labeyrie's idea to retrieve diffraction limited images rather than just the speckle transform of the stars. However, the result of their processing produces an image no better than a long time exposure photograph taken through a turbulent atmosphere. This mistake has been clearly shown by Miller et al. (1974).

Although Gerazi et al. (1972) measure the speckle transform of double stars, it was left to Bates et al. (1973) to point out that under certain conditions, Labeyrie's process could image some stars rather than just give their speckle transforms. Noting that the FT of the speckle transform is the autocorrelation of the star, Bates et al. (1973) recognised that the speckle transforms of the double stars produced by Gezari et al. could be compared to a FT hologram of two unresolvable coherent point sources. One point source provides a reference wavefront to the other point source. Confusion

often arises here since two point sources need to be coherent to form a hologram and stars are incoherent radiators. As an explanation, consider two quasi-monochromatic incoherent point sources radiating through an undistorted atmosphere and being imaged by a telescope. The distribution of light in the image plane comprises two bright spots with each spot approximating the intensity psf of the telescope. The power spectrum of this image is identical to a FT hologram of the original two point sources had they been radiating coherently. In the same way, when Labeyrie's technique is applied to two quasi-monochromatic incoherent point sources radiating through a turbulent atmosphere, the speckle transform derived from a number of distorted speckle images approximates a FT hologram had the two points been radiating coherently.

Bates et al. (1973) extend the comparison to a FT hologram and replace one of the two point sources by an extended object. The extended object may be large and continuous or composed of a group of individually unresolvable point sources. The speckle transform derived from the speckle interferograms of an extended object and the remaining point source now becomes a FT hologram of the extended object with the remaining object point source acting as a reference to the extended object. They call this speckle transform a speckle hologram and as with a true FT hologram, the image of the extended object is reconstructed in the rear focal plane of an optical FT system. A disadvantage is that in the image plane of the speckle hologram, as well as the true image of the extended object, there is its mirror image and both of these are superimposed on the autocorrelation of the extended object.

5.2 BASIC THEORY OF SPECKLE HOLOGRAPHY

Although Labeyrie's (1970) original treatment is improved by Dainty (1973) and Korff et al. (1973), the development here follows Labeyrie's simpler approach.

If a spatially incoherent, quasi-monochromatic image is formed through a grossly distorted imaging system with a limited aperture, the image intensity i(x,y) can be written as a convolution of the best available image intensity c(x,y) and the psf of the distortion h(x,y);

$$i(x,y) = c(x,y) \otimes |h(x,y)|^2 \tag{5.2}$$

It is more convenient to define $|h(x,y)|^2$ as the intensity psf, since $|h(x,y)|^2$ is the intensity of the image of a point source radiating in the object plane. Let

$$d(x,y) \equiv |h(x,y)|^2 \tag{5.3}$$

so that

$$i(x,y) = c(x,y) \otimes d(x,y) \tag{5.4}$$

which is the general equation for any distorted, spatially incoherent, quasi-monochromatic imaging system.

Consider an optical system imaging a star through the earth's atmosphere. The nearest star is at least 10^{12} miles distant, and a major portion of the earth's atmosphere lies within 10^{1} miles of the earth's surface.

Thus the phase distortion can be assumed to be introduced in a very narrow region just in front of the optical aperture.

Consider a point source in the object plane radiating through the atmosphere and being imaged by a telescope. If the exposure of the film in the image plane is much less than the redistribution time, then the intensity of the image may be defined as the instantaneous distortion $d_m(x,y)$. This image is composed of a number of randomly spaced speckles and each speckle is the approximate size of the diffraction limited disc. There is a certain region in the object plane within which the point source may move around and yet still produce the same intensity distribution in the image plane for the same distortion conditions. This is the isoplanatic region. If an extended object exists within this isoplanatic region then, if its best available image is $c(x,y)$, we may write the instantaneous image intensity $i_m(x,y)$ as a convolution

$$i_m(x,y) = c(x,y) \otimes d_m(x,y) \tag{5.5}$$

Labeyrie calls $i_m(x,y)$ a speckle interferogram. The speckle pattern taken at an instant in time does not resemble the speckle pattern taken at a later instant in time.

Let us first consider conventional astronomical photography. A long time exposure photograph is equivalent to averaging a large number of speckle interferograms. This may be written as

$$\sum_{m=1}^{N} i_m(x,y) = c(x,y) \otimes \sum_{m=1}^{N} d_m(x,y) \tag{5.6}$$

If N is sufficiently large, then the summation over N is almost equivalent to the ensemble average. That is, the

sample N is large enough so that statistics drawn from this sample are very close to the statistics drawn from the total population. Eqn (5.6) can now be written

$$\langle i_m(x,y)\rangle = c(x,y) \otimes \langle d_m(x,y)\rangle$$

where $\langle\ \rangle$ is the ensemble (total population) average. However, if all of c(x,y) lies within the isoplanatic region, the extent of c(x,y) is smaller than $\langle d_m(x,y)\rangle$. Therefore, eqn (5.6) may be written as

$$\langle i_m(x,y)\rangle \doteqdot \langle d_m(x,y)\rangle \tag{5.7}$$

The ensemble $\langle d_m(x,y)\rangle$ is the mathematical expression for the seeing disc and it can be seen from eqn (5.6) and eqn (5.7) that information larger than the resolution limit but smaller than the seeing disc is lost by conventional astronomical photography.

Consider the m^{th} speckle interferogram $i_m(x,y)$. It has a FT $I_m(\alpha,\beta)$ given by

$$I_m(\alpha,\beta) = C(\alpha,\beta)\ D_m(\alpha,\beta) \tag{5.8}$$

$D_m(\alpha,\beta)$ is a complex valued function whose phase is random above a spatial frequency of $1/\$$ where $\$$ is the diameter of the seeing disc. Therefore

$$\langle D_m(\alpha,\beta)\rangle = 0 \tag{5.9}$$

for spatial frequencies above $1/\$$. The expression for the long time exposure photograph in eqn (5.8) may also be given

in terms of its FTs,

$$\langle I_m(\alpha,\beta)\rangle = C(\alpha,\beta)\langle D_m(\alpha,\beta)\rangle \tag{5.10}$$

The effect of eqn (5.9) on eqn (5.10) is for $\langle D_m(\alpha,\beta)\rangle$ to act as a low pass spatial filter for frequencies below $1/\$$ and so the high frequency content of $C(\alpha,\beta)$ does not exist in $\langle I_m(\alpha,\beta)\rangle$. Consequently, detail in $c(x,y)$ that is smaller than the seeing disc does not exist in $\langle i_m(x,y)\rangle$. We now introduce speckle interferometry.

Consider a new function $E(\alpha,\beta)$ defined by

$$E(\alpha,\beta) = \langle |D_m(\alpha,\beta)|^2 \rangle \tag{5.11}$$

This is a smooth slowly varying function that exists out to $1/R$ where R is the linear dimension of the smallest detail that the imaging system can resolve. Consider also a function defined by

$$V(\alpha,\beta) = \langle |I_m(\alpha,\beta)|^2 \rangle \tag{5.12}$$

Combining eqn (5.11) and eqn (5.12) we get

$$V(\alpha,\beta) = |C(\alpha,\beta)|^2 \, E(\alpha,\beta) \tag{5.13}$$

Now let the radiating incoherent object be replaced by a single point radiator $\underline{c}(x,y)$ that is unresolvable by the undistorted imaging system. The underlined values denote distributions pertaining to the unresolvable object, so a

speckle interferogram of this object is

$$\underline{i}_m(x,y) = \underline{c}(x,y) \otimes d_m(x,y) \tag{5.14}$$

The ensemble average of the squared modulus of the FTs of these interferograms is

$$\begin{aligned} \underline{V}(\alpha,\beta) &= |\underline{C}(\alpha,\beta)|^2 \, \underline{E}(\alpha,\beta) \\ &= \underline{E}(\alpha,\beta) \end{aligned} \tag{5.15}$$

since $|\underline{C}(\alpha,\beta)|$ is necessarily constant where $\underline{E}(\alpha,\beta)$ has sufficient value. If a sufficient number of power spectra are averaged, then

$$\underline{E}(\alpha,\beta) = E(\alpha,\beta) \tag{5.16}$$

This equation implies that the statistics of the atmosphere must remain stationary over a period of time sufficiently long to record all the $i_m(x,y)$ and $\underline{i}_m(x,y)$. In general, it is sufficient to divide $V(\alpha,\beta)$ determined from eqn (5.13) by $\underline{V}(\alpha,\beta)$, determined from eqn (5.15), to obtain a reasonable estimate of $|C(\alpha,\beta)|^2$. Denote the result of this operation by $\hat{V}(\alpha,\beta)$, then

$$\begin{aligned} \hat{V}(\alpha,\beta) &= V(\alpha,\beta)/\underline{V}(\alpha,\beta) \\ &= |C(\alpha,\beta)|^2 \, E(\alpha,\beta)/\underline{E}(\alpha,\beta) \end{aligned} \tag{5.17}$$

Assuming that the statistics of the atmosphere remain stationary, we have that

$$\hat{V}(\alpha,\beta) = |C(\alpha,\beta)|^2 \tag{5.18}$$

The distribution $\hat{V}(\alpha,\beta)$ is called a speckle transform. As the speckle transform exists out to 1/R, information up to the diffraction limit of the telescope is contained in $\hat{V}(\alpha,\beta)$. The FT of $\hat{V}(\alpha,\beta)$ produces the autocorrelation of the object. If $\hat{v}(x,y)$ is the FT of $\hat{V}(\alpha,\beta)$ then

$$\hat{v}(x,y) = c(x,y) * c(x,y) \tag{5.18}$$

Some information regarding the object can be found from $\hat{v}(x,y)$, however, this may be ambiguous unless it is known that the object, and thus $c(x,y)$, is centrosymmetric.

5.2.1 Speckle Holography

There is a certain class of asymmetric objects that permit the reconstruction of images from the speckle transform. The essential conditions are that the asymmetric object should consist of at least two distinct parts and that the speckle transform have the character of a hologram.

Refer now to Fig. 5.1. Consider an object consisting of two separate parts. Then $c(x,y)$, the best available undistorted image also consists of two parts.

$$c(x,y) \equiv c_a(x,y) + c_b(x,y) \tag{5.19}$$

where A and B are the extents of each part along their bisecting axis and S is the separation between their nearest parts.

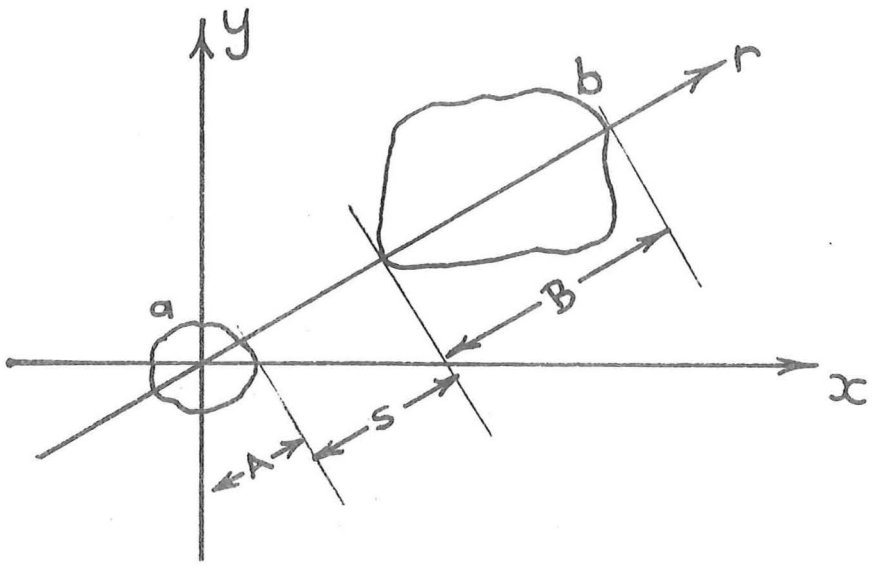


Fig. 5.1 An object consisting of two separate parts

a) $c_a(x,y)$ b) $c_b(x,y)$

The FT of the speckle transform $\hat{v}(\alpha,\beta)$, produced by Labeyrie's process now separates into four parts.

$$\begin{aligned}\hat{v}(x,y) &= c(x,y) * c(x,y) \\ &= c_a(x,y) * c_a(x,y) + c_b(x,y) * c_b(x,y) \\ &\quad + c_a(x,y) * c_b(x,y) + c_a(x,y) \otimes c_b(x,y) \qquad (5.20)\end{aligned}$$

If $S \geq L$ (5.21)

where L is the greater of A or B, then $c_a(x,y) * c_b(x,y)$ is completely separated from the other three terms. This is shown in Fig. 5.2.

(It is interesting to compare eqn (5.20) with the image distribution reconstructed from a FT hologram in section 2.2.1

$$\begin{aligned}u(x,y) &= r(x,y) * r(x,y) + o(x,y) * o(x,y) \\ &\quad + r(x,y) * o(x,y) + r(x,y) \otimes o(x,y) \qquad (5.21)\end{aligned}$$

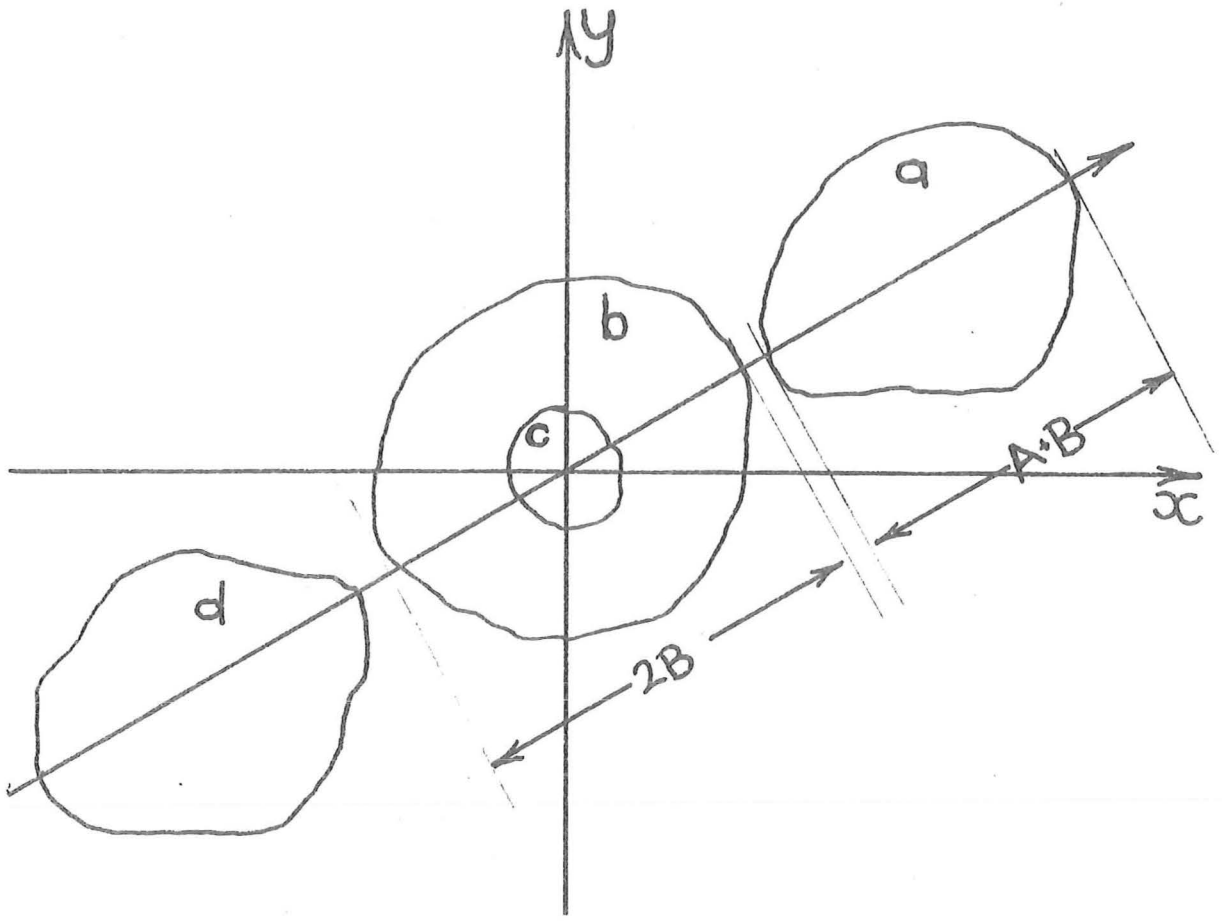


Fig. 5.2 The Autocorrelation of a two part object.

a) the cross-correlation $c_a(x,y) * c_b(x,y)$

b) the autocorrelation $c_b(x,y) * c_b(x,y)$

c) the autocorrelation $c_a(x,y) * c_a(x,y)$

d) the convolution $c_a(x,y) \otimes c_b(x,y)$.

where $r(x,y)$ is a point source radiator in the object plane coherent with an extended object, $o(x,y)$.)

Define the cross-correlation in eqn (5.19) as $\hat{v}_t(x,y)$, given by

$$\hat{v}_t(x,y) = c_a(x,y) * c_b(x,y) \tag{5.22}$$

If $c_a(x,y)$ is so small that it is unresolved by the optical system we may replace $c_a(x,y)$ by a two-dimensional

delta function. So

$$\hat{v}_t(x,y) = \delta(x,y) * c_b(x,y)$$

$$= c_b(x,y) \qquad (5.23)$$

Therefore, the FT distribution of a speckle hologram, provided one part is smaller than the diffraction limit, contains the image of the larger part of the object.

5.3 EXPERIMENTAL PROCEDURE

We decided to verify the theory of speckle holography (developed in the previous section) by simulating an astronomical imaging system on an optical bench.

The imaging system consists of a lens, with a focal length of 330 mm, which has an aperture stop of 5 mm in diameter. A sheet of shower glass was placed just in front of the lens on the side that was towards the object. Since the shower glass has a non-uniform thickness, it introduces phase distortion into the imaging system. This is equivalent to the phase distortion that is introduced instantaneously by the atmosphere to an astronomical imaging system. As a series of speckle interferograms was photographed, the shower glass was randomly moved between each exposure to simulate the change in the atmosphere.

The light radiating from a star has a range of optical frequencies and it is spatially incoherent across the surface of the star. Since only a narrow band of frequencies is required for the process, a 3.5 mW 6328 Å laser was used to simulate the stellar source. Spatial incoherence was obtained

by making the light leaving the laser strike a rapidly moving diffuser, so that the light behind this diffuser was spatially incoherent and quasi-monochromatic. A thin copper sheet was placed behind the moving diffuser in the object plane of the imaging lens. The perimeter of the object to be imaged was cut out from this opaque sheet. Thus, across the object plane, this illuminated plate simulated a quasi-monochromatic, spatially incoherent stellar source. The geometrical details of this optical simulation are illustrated in Fig. 5.3.

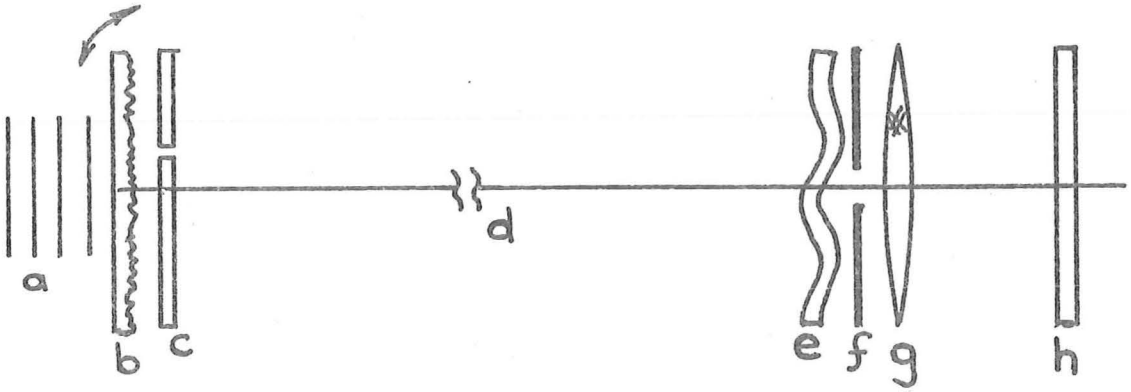


Fig. 5.3 The simulated optical system

a) Collimated coherent wavefronts
b) Moving diffuser
c) Object plane
d) Non-distorting medium
e) Distorting medium
f) 5 mm aperture stop
g) Lens
h) Image and film plane

We started by copying Labeyrie's process. The first object that we examined was an unresolvable point source. Consequently, a small hole was drilled in the copper sheet in the object plane so that the 5 mm pupil aperture of the imaging lens was just unable to resolve it. (This was checked by measuring the size of the coherent radiation pattern of the small hole. The diameter of the main lobe of the Airy disc

was just greater than the 5 mm aperture stop in the plane of the lens.)

A sheet of Kodak Panatomic X was pre-exposed to an even diffuse light for 1/125 seconds. Film emulsion has a certain amount of inertia, and it requires a pre-exposure to activate the film so that a subsequent exposure on the same piece of film is linear. The shower glass was placed in a random transverse position, and the film was exposed to the light from the object for two minutes. The film was advanced to the next pre-exposed frame, and with the shower glass in a different position, a second image was recorded. This procedure was continued until 20 images were photographed. Only 20 images were recorded since this was the largest number of photographs that could be developed at one time.

All the images recorded on the film were negatives. That is, the most intense portions of the image distribution were the least transmissive portions of the developed film. Each negative transparency was then reversed onto another film to make a positive transparency that linearly related to the original image intensity. Each negative was reversed onto Agfa Copex developed so that the overall gamma of both photographic steps is 2.0. Appendix A deals with the linearity requirements of photographic film.

The positive transparencies are the speckle interferograms. After all the speckle interferograms had been recorded a coherent FT system was set up and a sheet of pre-exposed Agfa Copex Pan Rapid film was placed in the rear focal plane of the FT lens. The first interferogram $i_1(x,y)$ was placed in a liquid gate (Appendix C) to remove the effect of the film

having a variable thickness and the film-liquid gate combination inserted in the front focal plane of the FT lens. The Copex film was exposed to the FT of $i_1(x,y)$ for 1/500 sec. A characteristic of the pre-exposed film emulsion ensures that the latent image in the undeveloped film records the two-dimensional power spectrum of $i_1(x,y)$.

The sheet of Copex film was left in place and the first interferogram was replaced by the second interferogram $i_2(x,y)$. The Copex film was then exposed for 1/500 sec. to the power spectrum of the second interferogram.

This procedure was repeated for all the twenty interferograms. The Copex film was then removed from the optical bench and developed for 9 mins in Rodinol at a 1/25 strength to a gamma of two. The developed film is the speckle transform of the object distribution used to make the speckle interferograms.

5.4 RESULTS

Fig. 5.4 illustrates the relevant features of the single unresolvable point source. 5.4(a) shows the best available image of a single point source object. Fig. 5.4(b) illustrates the effect of including the phase distorting plate. The image is broken up into a myriad of tiny randomly scattered speckles, spread over a very much larger area. This speckle pattern shifts and changes with the movement in the atmosphere. A normal long time exposure was made of this changing pattern and its resulting image, which is the seeing disc, is shown in 5.4(c). This is the type of image produced by conventional astronomical imaging procedures. It can be

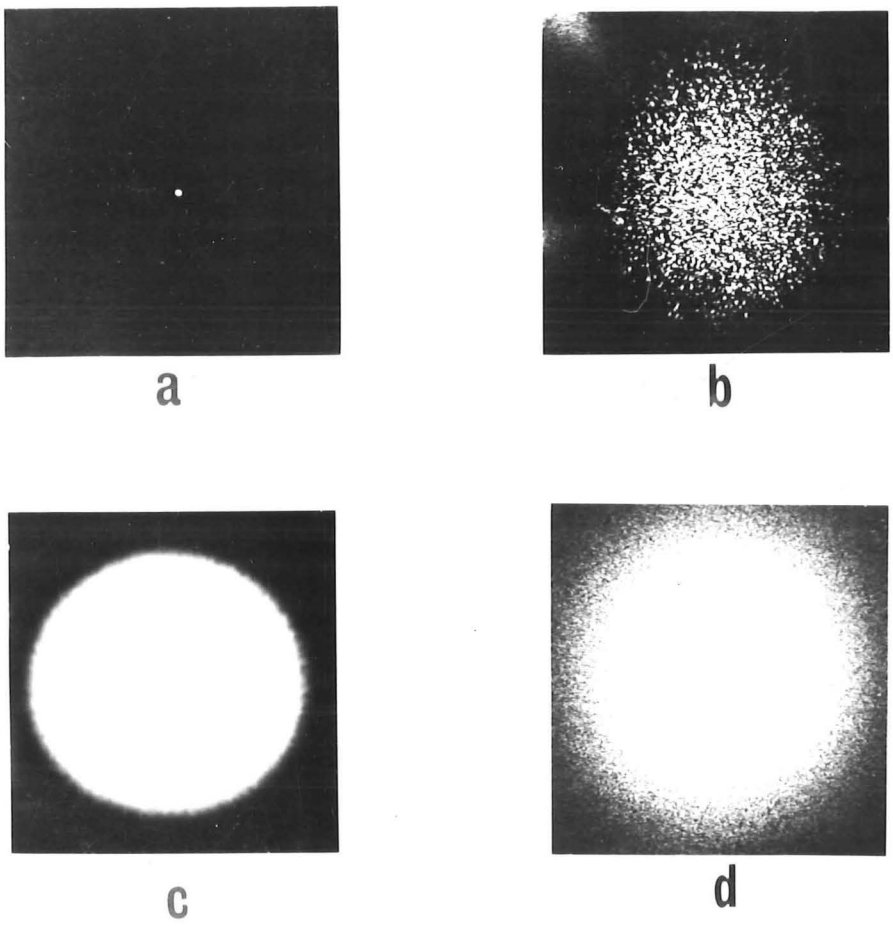


Fig. 5.4 The features of a single unresolvable object.

(a) The best available image formed when the phase distorting plate is removed.
(b) A speckle interferogram.
(c) A conventional long time exposure, the time average of the speckle interferograms.
(d) The speckle transform.

seen that the exact position of the star is quite uncertain. The speckle transform of the point source, that is $\underline{V}(\alpha,\beta)$, is shown in Fig. 5.4(d).

A second object was placed in the object plane which was a group of four 'stars'. Each 'star' was only just unresolvable by the undistorted imaging system. Thus, each hole drilled in the copper plate was the same size as the hole used for the single point. The holes were arranged so that one star was separated from the remaining stars. The intensity of this star was three times as great as each of the remaining stars. The separation used the minimum required by eqn (5.21). The experimental procedure was repeated as for the single star.

Fig. 5.5 illustrates the relevant features of this object. Fig. 5.5(a) shows the best available image that could be obtained, while 5.5(b) shows $i_m(x,y)$ and demonstrates the effect of the phase distorting plate. Note that because four point sources pass approximately six times more light than the single point source, the speckle interferogram 5.5(b) has had a greater exposure than Fig. 5.4(b). However, the conventional long time exposure of the four star object, Fig. 5.5(c), shows little difference from the long time exposure of the single star, Fig. 5.4(c). This demonstrates the amount of detail that is being lost when long time exposures are used to image stellar objects.

The procedure to produce the speckle transform from the speckle interferograms is the same for the multiple star as for the single star. Each speckle interferogram transparency is placed in the front focal plane of a Fourier transform lens and illuminated with spatially coherent light. The Fourier

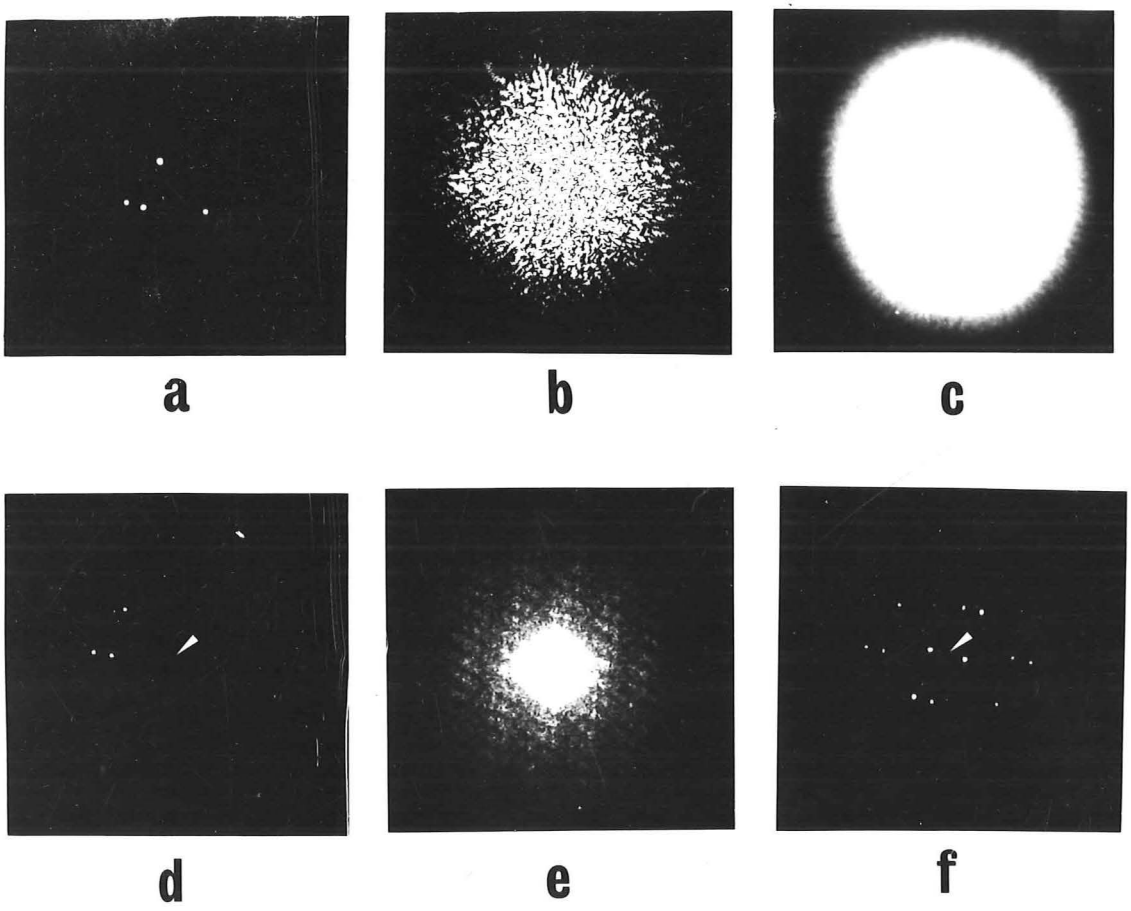


Fig. 5.5 Multiple Object.

(a) The best available image when the phase distorting plate is removed.
(b) A speckle interferogram.
(c) Time average of speckle interferograms.
(d) Reconstruction of object where the position of the reference is indicated by an arrowhead.
(e) Speckle hologram.
(f) Autocorrelation of the object where the position of the autocorrelation of the object is indicated by an arrowhead.

transform distribution of this transparency, located in the rear focal plane of the lens, is photographed, which automatically records the two-dimensional power spectrum. The speckle interferogram is then replaced by another interferogram and its power spectrum added to the power spectrum of the first interferogram. The result of the addition of the power spectra of all the interferograms is shown in Fig. 5.5(e). This is $V(\alpha,\beta)$. The speckle transform $\hat{V}(\alpha,\beta)$ is obtained by dividing $V(\alpha,\beta)$ by $\underline{V}(\alpha,\beta)$ shown in eqn (5.17). However, after a visual inspection of $V(\alpha,\beta)$ Fig. 5.5(e) and $\underline{V}(\alpha,\beta)$, Fig. 5.4(d), it was decided to omit this step, since the division would change $V(\alpha,\beta)$ very little, but would increase the effects of the noise. That is, by inspection, we determine that

$$\hat{V}(\alpha,\beta) \approx V(\alpha,\beta) \tag{5.24}$$

Note, that although the individual speckle interferograms of single star and groups of stars have much the same character, their relative speckle transforms $V(\alpha,\beta)$ and $\hat{V}(\alpha,\beta)$ are markedly different.

To complete the imaging process, the Fourier transform of $\hat{V}(\alpha,\beta)$ is performed optically, to give $\hat{v}(x,y)$. The result of this operation is illustrated in Fig. 5.5(f). There is usually a very intense spot of light in the centre of the two reconstructed images, that blots out the interesting portions of the images, so this is masked by a small opaque stop. The position of the exact centre is indicated in the figure by a small arrowhead. The relevant image $\hat{v}_t(x,y)$ in $\hat{v}(x,y)$ shown

in Fig. 5.5(f) and it is taken out and printed below the illustration of the best available image, so that a direct comparison can be made.

To summarise, the result of imaging a multiple star (Fig. 5.5(a)) using conventional long-time exposure, is shown in Fig. 5.5(c) and using speckle holography is shown in Fig. 5.5(d).

The final object used to test the experimental procedure is a larger asymmetrical 'star', a distance away from an unresolvable 'star'. As before, the power spectra of the individual speckle interferograms are added altogether on the one piece of film. This is $V(\alpha,\beta)$. It was decided that

$$\hat{V}(\alpha,\beta) \;\hat{\approx}\; V(\alpha,\beta) \tag{5.24}$$

So that $V(\alpha,\beta)$ illustrated in Fig. 5.6(d) approximates the desired speckle hologram. The transparency of $\hat{V}(\alpha,\beta)$ is Fourier transformed optically to give the autocorrelation $\hat{v}(x,y)$ and the desired image is shown in Fig. 5.6(c).

Some non-linearities in the process now become apparent. This is indicated by the halo around the larger star shown in Fig. 5.6(c). We consider that this is due to the loss of the central spatial frequencies in the speckle hologram.

The central portions of $\hat{V}(\alpha,\beta)$ shown in Fig. 5.5(e) and 5.6(d) are blank. This is where $\alpha,\beta \hat{\approx} 0$ and this is the part of the speckle hologram which determines the average level of illumination. If this central portion is not recorded faithfully, non-linearities, typically the halo effect, are seen, since the hologram acts as a high pass filter. The reason for this section of the transparency of $\hat{V}(\alpha,\beta)$ being blank is that the film is not able to record the high level of

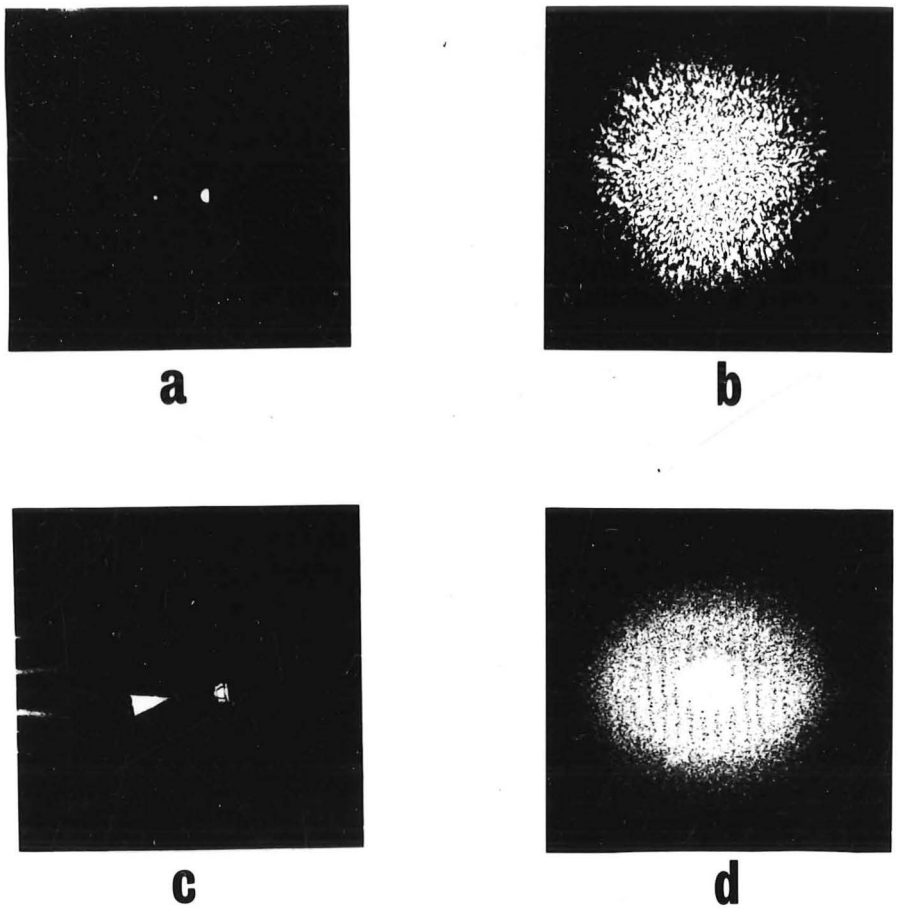


Fig. 5.6 Double Object.

(a) Best available image when the phase distorting plate is removed.
(b) A speckle interferogram.
(c) Reconstruction of object where the position of the reference object is indicated by an arrow.
(d) Speckle hologram.

incident illumination faithfully. Once the film has received enough light to produce its maximum density, no amount of further light will make any change. This is due to the limited density range of the film.

5.5 COMMENTS

A problem in the practical use of speckle holography is to determine how many speckle interferograms are needed to eventually produce a good visual image. If nothing is known about the relative sizes and intensities of the sources that constitute the object, then the best method is to photograph as many interferograms as can be handled conveniently.

If the relative intensities of the sources that constitute an object are known from some previous photometry measurements, then it is possible to obtain an estimate of the number of interferograms required.

A speckle interferogram is defined as

$$i_m(x,y) = c(x,y) \otimes d_m(x,y) \tag{5.25}$$

but for the purposes of this analysis we confine the mathematics to one dimension, so

$$i_m(x) = c(x) \otimes d_m(x) \tag{5.26}$$

Consider a point source radiating through the atmosphere. Across the lens aperture of a one-dimensional 'telescope' denoted by the coordinate x_ℓ, there is a complex distribution of light denoted by $H(x_\ell)$. The aperture of this 'lens' has an

effective 'diameter' ζ and is denoted by rect $(\frac{1}{\zeta} x_\ell)$ where rect $(\frac{1}{\zeta} x_\ell)$ is 1 inside $-\zeta/2$ to $\zeta/2$ and zero elsewhere. Thus $D_m(\alpha)$, the FT of the instantaneous image $d_m(x)$ is

$$D_m(\alpha) = \frac{4}{\pi\zeta^2} \int_{-\infty}^{\infty} H(x_\ell)\ H^*(x_\ell - \bar{\lambda} f\alpha)$$

$$\text{rect}\ (\tfrac{1}{\zeta} x_\ell)\ \text{rect}(\tfrac{1}{\zeta} x_\ell - \tfrac{1}{\zeta} \bar{\lambda} f\alpha)\ dx_\ell \qquad (5.27)$$

where $\bar{\lambda}$ is the mean wavelength and f is the focal length of the lens (Born and Wolf, 1959, p485). Now

$$D_m(\alpha) = 0 \qquad \alpha \geq \frac{\zeta}{\bar{\lambda} f} \qquad (5.28)$$

because the autocorrelation of the rect ( ) function is always zero above $\zeta/\bar{\lambda}f$. If there is no distortion in the atmosphere, $H(x_\ell)$ is smooth and usually constant, so that the spatial frequency distribution in the image is determined solely by the autocorrelation of the rect ( ) function. Consequently, $\zeta/\bar{\lambda}f$ is the <u>maximum</u> spatial frequency of a <u>distortionless</u> optical system. This is the resolution limit of the telescope.

In speckle interferometry, we are interested in spatial frequencies above 1/\$ where \$ is the angular extent of the seeing 'disc', and below $\zeta/\bar{\lambda}f$ the resolution limit. As well, the complex amplitude $H(x_\ell)$ exhibits modulus and phase fluctuations which become uncorrelated for a separation greater than r_s which is usually less than ζ (Fried, 1966). Consequently, Korff et al. (1972) simplify eqn (5.27) to give

$$D_m(\alpha) = \left(\frac{r_s}{\zeta}\right)^2 \sum_{n=1}^{N} H(\xi_n)\ H^*(\xi_n - \bar{\lambda} f \alpha) \qquad (5.29)$$

for $1/\$ < \alpha \leqslant \zeta/\bar{\lambda}f$.

From eqn (5.29), $D_m(\alpha)$ for any spatial frequency in this range is the sum of N independent, random vectors, and if N is sufficiently large (> 30), then regardless of the distribution of $H(x_\ell)$, $D_m(\alpha)$ has a Rayleigh distribution in amplitude and a uniform distribution in phase. As the phase is distributed uniformly

$$\langle D_m(\alpha) \rangle = 0 \qquad (5.30)$$

for $1/\$ < \alpha \leqslant \zeta/\bar{\lambda}f$, and information above the limit set by the seeing disc is lost. As the modulus $|D_m(\alpha)|$ is a Rayleigh distribution, then the square of the modulus $|D_m(\alpha)|^2$ is an exponential distribution.

Now consider the effect of adding a number of independent $|D_m(\alpha)|^2$. By considering an exponential distribution, a χ^2 distribution with two degrees of freedom, and using characteristic functions we can show the probability density function resulting from the addition of M power spectra is a χ^2 probability density function with 2M degrees of freedom (Fraser, 1958). In fact, if M is large enough ($\geqslant 30$), then this distribution approximates a normal distribution centred at 2M with a standard deviation $2M^{\frac{1}{2}}$.

Define the signal to noise ratio (S/N) of the speckle transform derived from M interferograms as the mean of the distribution divided by the standard deviation,

$$(S/N)_M = M^{\frac{1}{2}} \qquad (5.32)$$

A speckle hologram derived from M speckle interferograms is a grating of fringes embedded in random noise, and what we would like to know is the number M of speckle interferograms required before the fringes become visible. This closely corresponds to retrieving images immersed in random noise, but in this case our 'image' is the fringes in the speckle hologram.

Consider a multiple object with a best available image c(x) consisting of two incoherent point sources. Then

$$|C(\alpha)|^2 \propto 1 + \frac{i_0}{i_1} \cos\left(\frac{2}{S}\alpha\right) \qquad (5.33)$$

where i_0/i_1 is the relative radiating intensities of the two point sources, and S is their separation in the object plane. i_0/i_1 is often called the modulation depth of the fringes. If i_0/i_1 is 1/3 then the $(S/N)_M$ ratio necessary in the speckle hologram before the fringes in $|C(\alpha)|^2$ become visible is 3. Consequently from eqn (5.32) it is necessary to add nine interferograms.

If i_0/i_1 is 1/5 the modulation depth of the fringes is less, so it is necessary to add 25 interferograms for the fringes to become visible.

Consider now a speckle hologram formed from a radiating object c(x) consisting of two parts: an extended object f(x) and a point source separated from the centre of f(x) by S. In this case the modulation depth of the fringes varies across the speckle hologram, so that different parts of the hologram will require different $(S/N)_M$ for the fringes to be visible there. Consequently, the $(S/N)_M$ is determined by considering

the modulation depth of the fringes around the highest spatial frequencies that are required. For instance, $|C(\alpha)|^2$ of the two part object is

$$|C(\alpha)|^2 \propto 1 + F(\alpha) \cos(\frac{2}{S} \alpha) \tag{5.34}$$

We are interested in all spatial frequencies up to α_1 and the modulation depth of the fringes at α_1 is $F(\alpha_1)$ so that

$$M \geq [1/F(\alpha_1)]^2 \tag{5.35}$$

for the fringes around α_1 to be visible.

5.6 CONCLUSIONS

We have shown that a process developed to produce the autocorrelation of stars photographed through the atmosphere may be extended so that under certain conditions the process may produce images of stars. The essential condition is that there is an unresolvable bright star close to the unknown star we wish to image. Consequently, it may be worthwhile to tabulate the number of occasions that near occultations of stellar with solar objects occur. There is also the possibility of using occultations of stellar and solar objects with slow moving satellites.

CHAPTER 6

COMMENTS ON SUGGESTED IMPROVEMENTS TO SPECKLE INTERFEROMETRY

6.1 INTRODUCTION

Since the work reported in chapter 5 has been completed, a suggestion has been advanced (Knox and Thompson, 1974) which claims that it is not necessary to use the holographic approach to determine the undistorted image of a star(s). Encouraging results of computer simulations are reported by the authors. The first part of this chapter explains their proposed technique and then outlines some practical details that must be considered before it can be used in practice.

The second part of this chapter discusses an extension to speckle interferometry suggested by Boidin and Labeyrie (1974) to handle very faint stars.

6.2 KNOX-THOMPSON PROCESSING

Knox and Thompson (1974) have made the remarkable suggestion that the phase as well as the modulus of the FT of the object may be recoverable from a set of speckle interferograms. Recall that the basic equation describing Labeyrie's process, in one dimension, is

$$\langle |I_m(\alpha)|^2 \rangle = |C(\alpha)|^2 \langle |D_m(\alpha)|^2 \rangle . \qquad (6.1)$$

$|C(\alpha)|^2$ is determined by dividing $\langle |I_m(\alpha)|^2 \rangle$ by $\langle |D_m(\alpha)|^2 \rangle$.

Knox and Thompson base their argument on the relation

$$\langle I_m(\alpha) I_m(\alpha+\Delta)\rangle = C(\alpha)C^*(\alpha+\Delta)\langle D_m(\alpha) D^*(\alpha+\Delta)\rangle \qquad (6.2)$$

where Δ is a fixed small displacement. Note that when

$$\Delta = 0$$

eqn (6.2) reduces to eqn (6.1).

We may define a new function

$$L(\alpha) \equiv \frac{\langle I_m(\alpha) I_m^*(\alpha+\Delta)\rangle}{\langle D_m(\alpha) D_m^*(\alpha+\Delta)\rangle} \qquad (6.3)$$

For $L(\alpha)$ to be useful, the denominator of eqn (6.3) must be greater than the random noise for all spatial frequencies where the numerator is not zero. So that, if R is the linear dimension of the smallest resolvable detail, then $\langle D_m(\alpha) D_m(\alpha+\Delta)\rangle$ must be significant out to a spatial frequency of 1/R. It follows from eqn (6.2) and (6.3) that

$$L(\alpha) = C(\alpha)C^*(\alpha+\Delta) \qquad (6.4)$$

We may define both $L(\alpha)$ and $C(\alpha)$ in terms of modulus and phase

$$L(\alpha) \equiv |L(\alpha)|\ e^{j \text{ phase } L(\alpha)} \qquad (6.5)$$

and

$$C(\alpha) \equiv |C(\alpha)|\ e^{j \text{ phase } C(\alpha)} \qquad (6.6)$$

Considering only the phase terms of eqn (6.4)

$$\text{phase } L(\alpha) = \text{phase } C(\alpha) - \text{phase } C(\alpha+\Delta) \qquad (6.7)$$

Now phase $C(0) = 0$ (6.8)

because $C(x)$ is real. When

$$\alpha = 0$$

$$\text{phase } C(\Delta) = -\text{phase } L(0) \tag{6.9}$$

It follows from eqn (6.7) and (6.9) that

$$\text{phase } C(2\Delta) = \text{phase } C(\Delta) - \text{phase } L(\Delta)$$

or if n is any non-negative integer

$$\text{phase } C((n+1)\Delta) = \text{phase } C(n\Delta) - \text{phase } L(n\Delta) \tag{6.10}$$

Using eqn (6.10) we can calculate the phase of $C(\alpha)$ sampled at intervals of Δ. If W is the total width of the extended stellar object, then we require samples spaced by $1/W$ or less. If the samples are spaced by greater than $1/W$, then the phase of $C(\alpha)$ is undersampled and aliasing occurs. Thus the <u>maximum</u> allowable value of Δ is $1/W$.

Since $|C(\alpha)|$ can be determined from eqn (6.1) and the phase of $C(\alpha)$ can be determined from (6.10), $C(\alpha)$ can be determined uniquely from a series of speckle interferograms. Knox and Thompson (1974) carry out a computer simulation of this one-dimensional imaging and restoration process, achieving encouraging results.

The mathematics of this process may be easily extended into two dimensions. We first calculate $L(\alpha,\beta)$ for a small displacement Δ along the α axis. We can then calculate any phase of $C(\alpha,\beta)$ relative to the value at the β axis. In

effect the α,β plane is divided up into horizontal strips and any value along any one horizontal strip is known relative to the value where the strip crosses the β axis. The complete phase is calculated by now computing $L_V(\alpha,\beta)$ for a small vertical displacement Δ^+. This means that the phase of $C(\alpha,\beta)$ is known along vertical strips. Consequently, knowledge of the phase of $L(\alpha,\beta)$ for vertical and horizontal displacements uniquely determines the phase of $C(\alpha,\beta)$. Mathematically

$$L(\alpha,\beta) = C(\alpha,\beta)\, C^*(\alpha+\Delta,\beta)$$

$$L_V(\alpha,\beta) = C(\alpha,\beta)\, C^*(\alpha,\beta+\Delta^+) \qquad (6.12)$$

$$\begin{aligned}\text{phase } C((n+1)\Delta,(m+1)\Delta^+) = {} & \text{phase } C(n\Delta,m\Delta^+) \\ & + \text{phase } C(0,m\Delta^+) - \text{phase } L(n\Delta,m\Delta^+) \\ & - \text{phase } L_V(0,m\Delta^+)\end{aligned} \qquad (6.13)$$

where n, m are non-negative integers.

6.2.1 Practical Considerations

In astronomical practice the i_m are images appearing successively in an image plane of a telescope as it tracks the object across the celestial sphere. The telescope is not always pointed at the same point on the celestial sphere. The angular error is called the boresight error. If there is no boresight error, the spatial frequencies in $L(\alpha,\beta)$ and $L_V(\alpha,\beta)$ are significant out to the resolution limit of the telescope. If the root-mean-square boresight error of the interferograms is greater than the smallest resolvable detail in the star, then the spatial frequencies in $L(\alpha,\beta)$ and $L_V(\alpha,\beta)$ are reduced near $1/R$.

Knox and Thompson propose to counteract the boresight errors by preprocessing the speckle interferograms before performing the FT. They assume that the centroid of the speckle interferogram remains constant from one speckle interferogram to the next. However, if the individual speckles in each speckle interferogram are randomly positioned then the centroid of the speckles is randomly positioned in each interferogram, having a mean position and a variance about this mean. It is possible that if there are enough speckles in each speckle interferogram, the variance of the centroid about its mean position may be small enough that the centroid can be assumed to be constant from one speckle interferogram to the next.

The success of the Knox-Thompson processing would seem to critically depend on the value of Δ. Since the phase of $C(\Delta)$ is a summation of $L(\Delta)$ (eqn 6.10) any errors in $L(\Delta)$ have a cumulative effect on $C(\Delta)$. Consequently the phase of $C(\alpha)$ for large α becomes inaccurate. Denote the error in $L(\alpha)$ by $E(\alpha)$, so that eqn (6.10) may be rewritten as

$$\text{phase } C((n+1)\Delta) = \text{phase } C(n\Delta) + \text{phase } L(n\Delta) + \sum_{m=0}^{n} \text{phase } E(m\Delta) \qquad (6.15)$$

As Δ is decreased, n is increased so that $n\Delta$ is a constant and the error in $C((n+1)\Delta)$ increases. Since the error is increased as Δ becomes smaller, it would seem optimal to make Δ as large as the sampling theorem allows. However, if there is a significant amount of boresight error, the sampling interval Δ must be small enough to remain unaffected by the boresight error. Let the mean-square-root boresight error be

ε, then the optimum sampling interval is

$$\Delta = \min \left(\frac{1}{W}, \frac{1}{\varepsilon}\right) \tag{6.16}$$

We now outline the procedure for calculating $C(\alpha,\beta)$ from a series of speckle interferograms taken of an extended stellar object and an unresolvable star. The statistics of the turbulence of the atmosphere are required to be stationary during the exposure of all the interferograms.

Labeyrie's process as outlined in Chapter 5 is first performed in order to determine $|C(\alpha,\beta)|$. This may be carried out optically or digitally although only optical processing seems to have been used to date. Knox-Thompson processing is performed next. The centroid of each interferogram is calculated and repositioned to the origin of the x,y plane. Now $\langle I_m(\alpha,\beta)\ I_m^*(\alpha+\Delta,\beta)\rangle$ is computed from the repositioned speckle interferograms of the extended stellar object, and $\langle D_m(\alpha,\beta)\ D_m^*(\alpha+\Delta,\beta)\rangle$ is computed from the repositioned speckle interferograms of the unresolvable star. Then $L(\alpha,\beta)$ as defined by eqn (6.3), is calculated by dividing $\langle I_m(\alpha,\beta)\ I_m^*(\alpha+\Delta,\beta)\rangle$ by $\langle D_m(\alpha,\beta) D_m(\alpha+\Delta,\beta)\rangle$. Next $L_V(\alpha,\beta)$ is calculated using an identical procedure. Finally the phase of $C(\alpha,\beta)$ is computed using eqn (6.13). The maximum value of n and m is N given by

$$N\Delta = 1/R$$

where R is the linear dimension of the smallest resolvable detail in the stellar object.

The actual computations in the Knox-Thompson process are done digitally since it is not possible to carry out the division to determine $L(\alpha,\beta)$, or to calculate the phase of $C(\alpha,\beta)$ by purely optical processing. Consequently, it would seem that the biggest drawback to this processing is the time taken to compute the phase of $C(\alpha,\beta)$. Each speckle interferogram is placed in the scanning aperture of the image scanner (Kennedy and Peters, 1974) and a sampled quantised approximation to $i_m(x,y)$ is stored in the computer. This scanning process is performed for each interferogram. Using the image scanner available it would take us 500 minutes to complete the scanning of twenty interferograms of a large star and twenty of an unresolvable star. There would also need to be a program to check that each interferogram was placed in exactly the same position as all the other interferograms. Lateral displacement in the scanning plane is equivalent to boresight error. It is possible that repositioning the centroid after the interferogram has been scanned and digitised may offset some of the error in lateral position in the image scanner.

There is the exciting possibility of direct on-line processing of speckle interferograms with a computer-backed image intensifier. This is where the image formed in the image plane of the telescope is fed directly into a computer, completely eliminating the photographic stage. By eliminating the photographic stage, the total time required for the Knox-Thompson processing could be much reduced.

The final computation in the combined Labeyrie plus Knox-Thompson process is to compute $c(x,y)$ from $C(\alpha,\beta)$, where $c(x,y)$ is the image of any extended stellar object.

6.4 THE FUTURE OF SPECKLE INTERFEROMETRY AND SPECKLE HOLOGRAPHY

The implications of the various experiments and theoretical evaluations of speckle interferometry are that errors introduced by the fluctuating atmosphere may be almost completely removed by processing a series of short exposure images. We may well find that higher resolution optical images may be produced by using ground based telescopes rather than orbiting space telescopes.

CHAPTER 7

CONCLUDING REMARKS

This thesis reports three useful applications of holography to image processing. Particular emphasis has been placed on the practicality of the techniques and in all cases details of experiments or experimental simulations have been reported.

7.1 SUGGESTIONS FOR FURTHER RESEARCH ON COMPUTER GENERATED HOLOGRAPHY

It would seem from the research reported in Chapters 3 and 4 that cgh's have definite limitations. Cgh's may successfully produce images of diffuse objects (Brown & Lohmann, 1969) and even of specular objects (as demonstrated in Chapter 3), however, it is not easy to get them to work successfully as inverse spatial filters. The reason for this latter failure is lack of resolution elements and lack of dynamic range in the cgh. Eventually, it may not be necessary to have such a large number of resolution elements (Lowenthal et al., 1974) and the lack of dynamic range will be the limiting factor. We have shown that there are two limitations on the attainable dynamic range. The first is the small difference in the reflectance of matt black paper and the reflectance of glossy white paper used to print out the computer artwork. A possible way to overcome this limitation is to imprint the hologram directly onto film. As more microphotograph computer memories become available, this direct imprinting

be feasible. However, the main limitation on all cghs, and on all holograms of specular objects, is the limited dynamic range of the photographic film used as the transparency. Until both the grain size of a film is decreased and the sensitivity increased, cgh's have limited usefulness as inverse filters.

7.2 SUGGESTION FOR FURTHER RESEARCH INTO SPECKLE HOLOGRAPHY AND RELATED TOPICS

Speckle holography has now been developed to the stage where it can be applied to real stellar objects. This requires an image intensifier coupled directly to an astronomical telescope, since very short exposures are required to ensure that the speckles do not move during the exposure.

The usefulness of the processing suggested by Knox and Thompson (1974) cannot yet be gauged properly. It is suggested that an optical simulation of this process should be carried out, in a similar fashion to the simulation reported in Chapter 5, before any attempt is made to process speckle interferograms from real stars.

Finally, it would seem that the application of speckle interferometry and its related techniques will soon produce clear undistorted images of stellar objects down to the resolution limit of any large aperture telescope.

APPENDIX A

PHOTOGRAPHIC FILM

A1 TYPES OF FILM

In all the optical experiments performed, photographic film was used to record the intensity distribution across the image plane. Two types of films were used: Kodak Panatomic-X and Agfa Copex Pan Rapid.

Panatomic-X is rated by the manufacturers (Kodak, 1969) at 40 ASA, and depending on the developer, it has the ability to resolve up to 80 lines per mm.

Copex is rated at 6 ASA (Agfa, 1968), so that this film requires more energy to form an image, than does Panatomic-X. Thus, it is termed a slower film. Copex is capable of resolving up to 500 lines per mm. Usually, the slower a film is rated, the greater is its resolving power.

A2 STRUCTURE OF THE PHOTOGRAPHIC FILM

Unexposed film consists of a sensitised emulsion on a transparent backing. The emulsion is a silver halide imbedded in gelatine. When sufficient light energy falls on the silver halide, it forms a latent image. This is not a visible image, but the silver halide that has received the energy is chemically activated. The film is then immersed in a liquid developing agent. This fluid reacts only with the activated silver halide and precipitates out the silver atoms.

The film is removed from the developer and immersed in a fixing agent. This fluid dissolves the silver halide that has not been activated. The film is washed and dried.

The image is now visible in a reversed or negative state. Where the greatest amount of light has fallen on the film, the emulsion is darkest, since this is where the greatest proportion of silver atoms exist . The silver atoms clump together to form microscopic grains.

A3 FILM PARAMETERS

Consider a film that is exposed to an even amount of illumination and then developed and fixed. This film is placed in an optical bench and illuminated with a uniform field of unit intensity i_0. If i_t is the intensity of the illumination passing through the film, then a measure of its ability to transmit light is the density d, given by

$$d = \log_{10} (i_0/i_t) \qquad (A.1)$$

There are many machines available that measure photographic density. The machine used in the optical experiments reported in this work is the Joyce scanning densitometer, made available by the Physics Department of the University of Canterbury.

Consider now an unexposed film that is illuminated by a non-uniform field. Denote the intensity across the field at the film plane by i(x,y). This field has an average intensity level, and the exposure of a film is determined by the average intensity multiplied by the length of time the shutter is

released. The units of exposure are energy per unit area, and are often given in Joules/cm^2.

To characterise a film it is necessary to know the change in density for the change in exposure. Now, film has a density range from 0.05 to approximately 2.5, and the upper limit of this range is determined by the type of film. The time that the shutter stays open is set so that the average exposure develops the film to its median density, say 1.25. Thus, exposures that are less than the average exposure give less density than 1.25, and exposures that are greater give a density greater than 1.25. If the density versus log exposure is plotted then there is a section around the median density that is linearly related to the log of the exposure. The slope of this linear region with respect to the log-exposure axis is defined as the gamma (γ) of the film. This is illustrated in the "Hurter-Driffield" curve in Fig. A1.

Gamma is given by

$$\gamma = \frac{\Delta d}{\Delta \text{ log exposure}} \qquad (A.2)$$

γ is a dimensionless number which usually lies between 0.55 and 4.0. However, the density is related to the amount of time the exposed film is immersed in the developer. So, Fig. A1 is the H-D curve for only one time of development. If three sheets of the same film are exposed to the same incident illumination and developed for different lengths of time, then there are three different H-D curves. This is shown in Fig. A2.

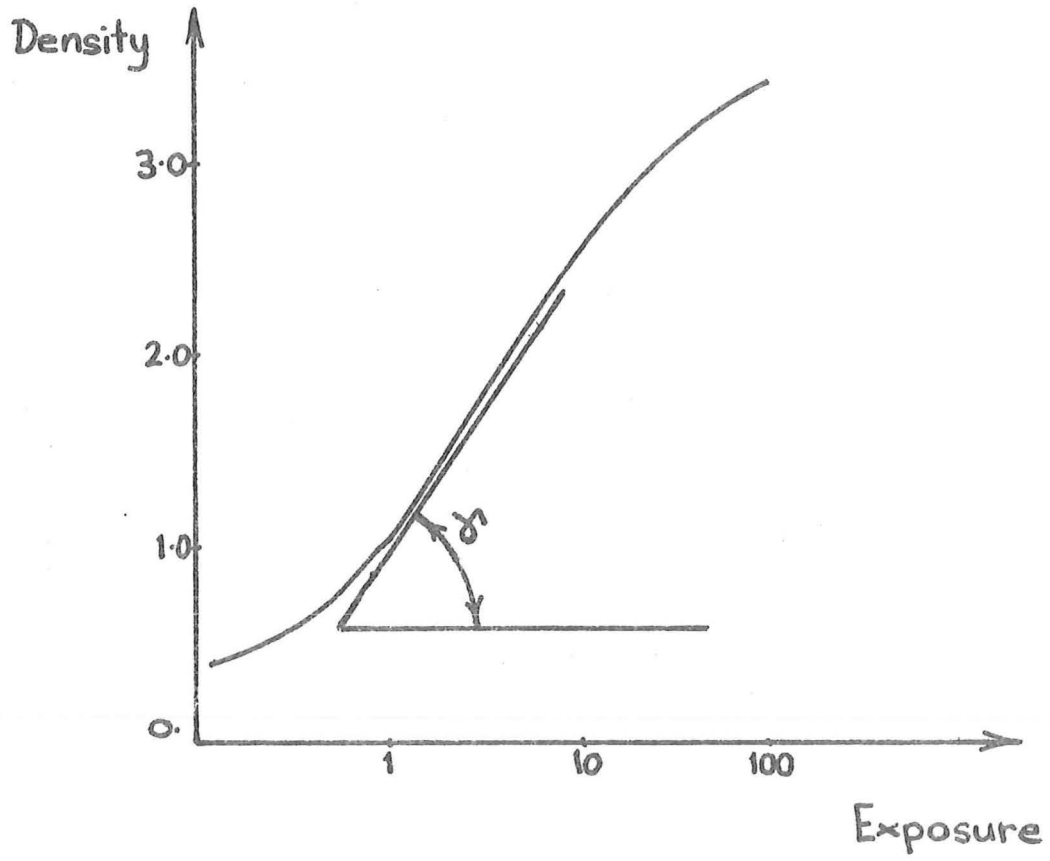


Fig. A1 H-D Curve for a Film

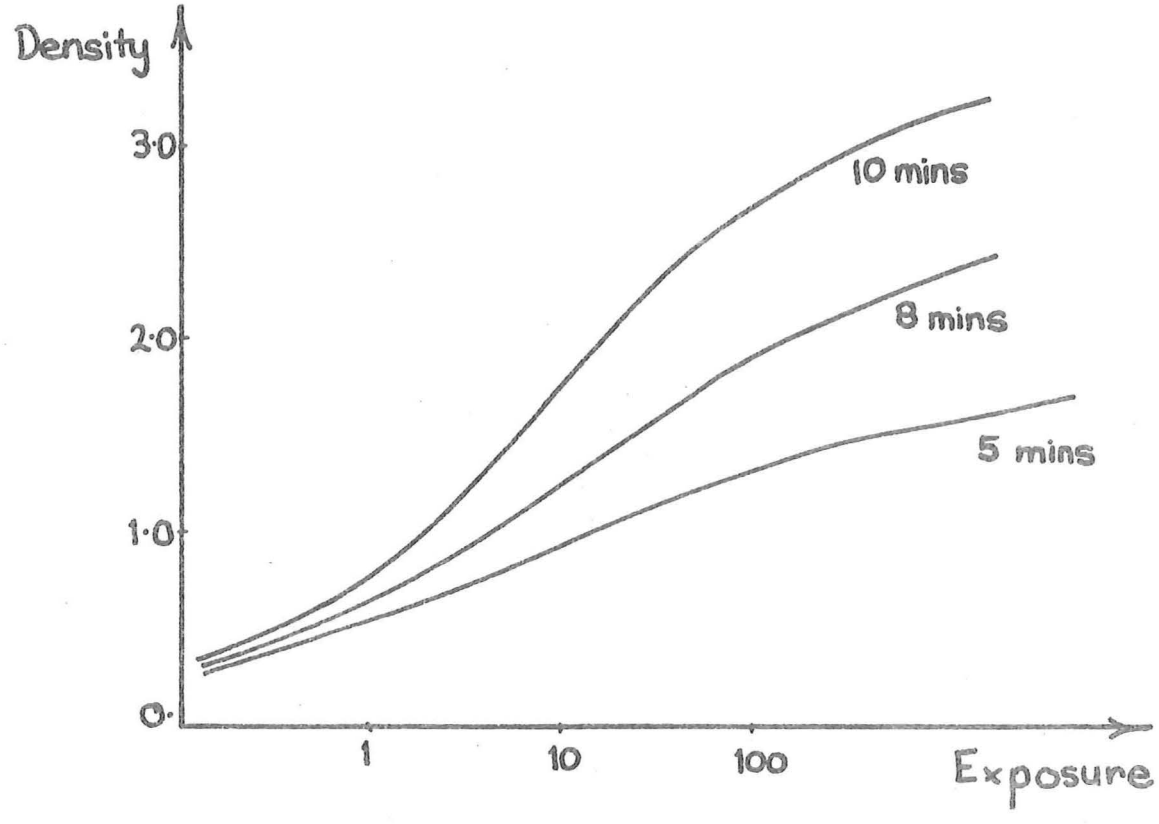


Fig. A2 H-D curves showing different times of development.

What is more useful to the researcher is a graph of the γ obtained by a certain time of development. Fig. A3 shows three curves for the two films used, at different strengths of developer. The developing agent is Agfa Rodinol.

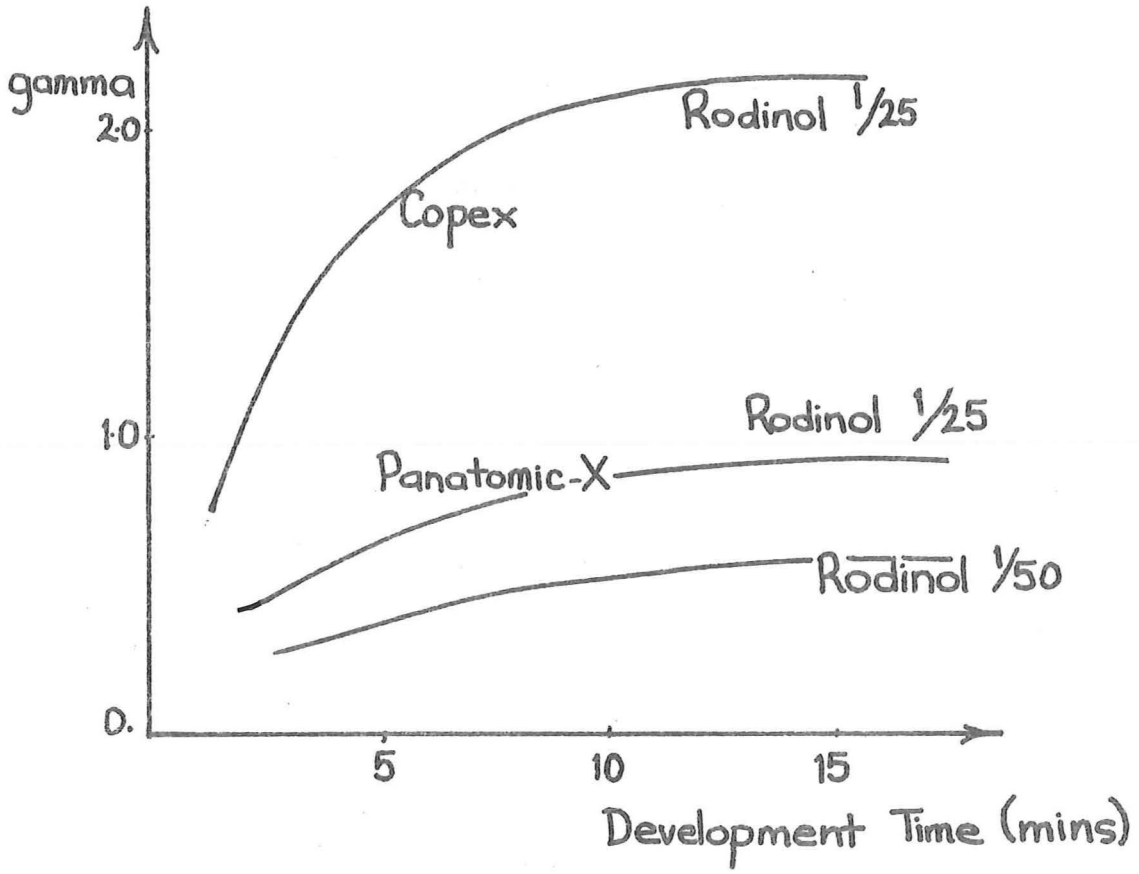


Fig. A3

A4 THE DEPENDENCE OF THE IMAGE ON GAMMA

If $i(x,y)$ is the intensity of the illumination incident on an emulsion and the film is developed to a certain γ then the developed film illuminated by a field of intensity $i_0(x,y)$ transmits an intensity distribution $i_1(x,y)$ where

$$i_1(x,y) = i_0(x,y)[i(x,y)]^{-\gamma} \qquad (A.3)$$

Normally the intensity of the light that illuminates the developed film is constant over the x,y plane, so that

$$i_1(x,y) = i_0[i(x,y)]^{-\gamma} \qquad (A.4)$$

These values are illustrated in Fig. A4.

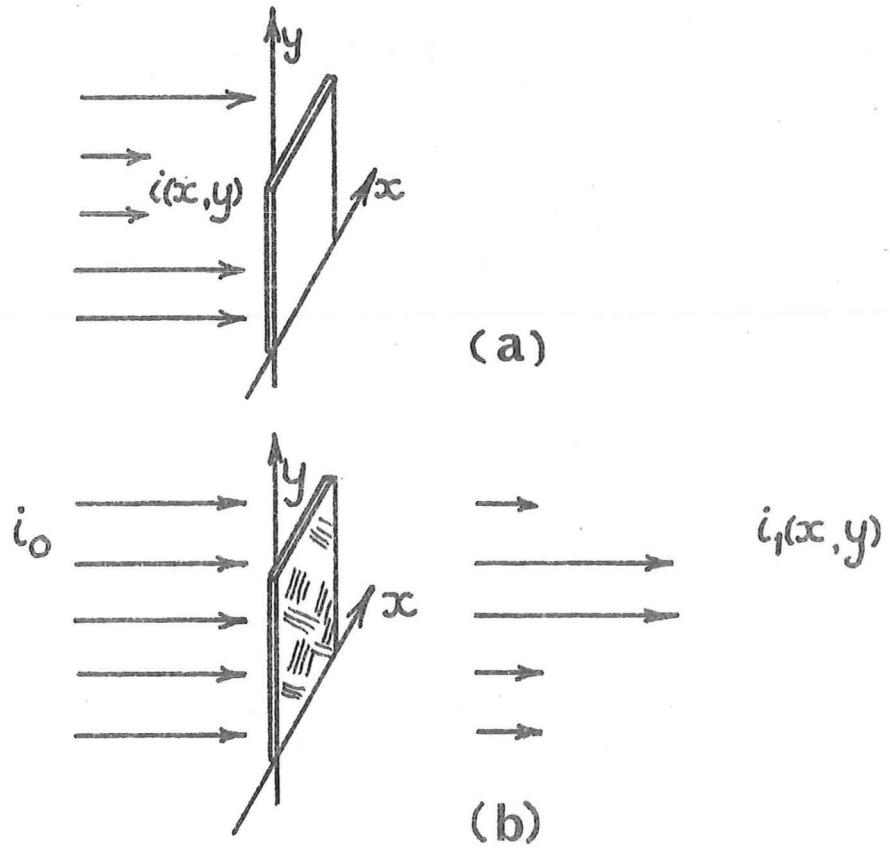


Fig. A4 Exposing a film.

a) Film exposed to incident intensity.

b) Developed negative exposed to unit intensity.

Since the developed film transmits the inverse of the intensity distribution that was incident on it, the developed film is usually called a negative.

In many cases it is required that a developed film transmits an intensity distribution that is proportional to the incident intensity, not its inverse. To obtain this the negative must be reversed, that is, a second negative must be made of the first negative. The second negative is often called a positive transparency. If the first negative is developed to a gamma of γ_N and the second negative developed to a gamma of γ_P, then the transmitted intensity distribution of the second negative (positive) is

$$i_2(x,y) = i_0^2[i(x,y)]^{-(\gamma_N\gamma_P)} \tag{A.5}$$

This is illustrated in Fig. A5.

The exact value of the overall gamma product $\gamma_N\gamma_P$ depends on the application that the positive is to be used for. If the positive is used in an incoherent system, then the transmitted intensity of the positive $i_2(x,y)$ is usually required to be exactly the same as the original incident intensity $i(x,y)$. Thus

$$\gamma_N\gamma_P = 1 \tag{A.6}$$

Any combination of individual γ may be used provided that no individual γ lies outside the permitted range for the film being used.

If the positive is subsequently used in a coherent optical system then

$$\gamma_N\gamma_P = 2.0. \tag{A.7}$$

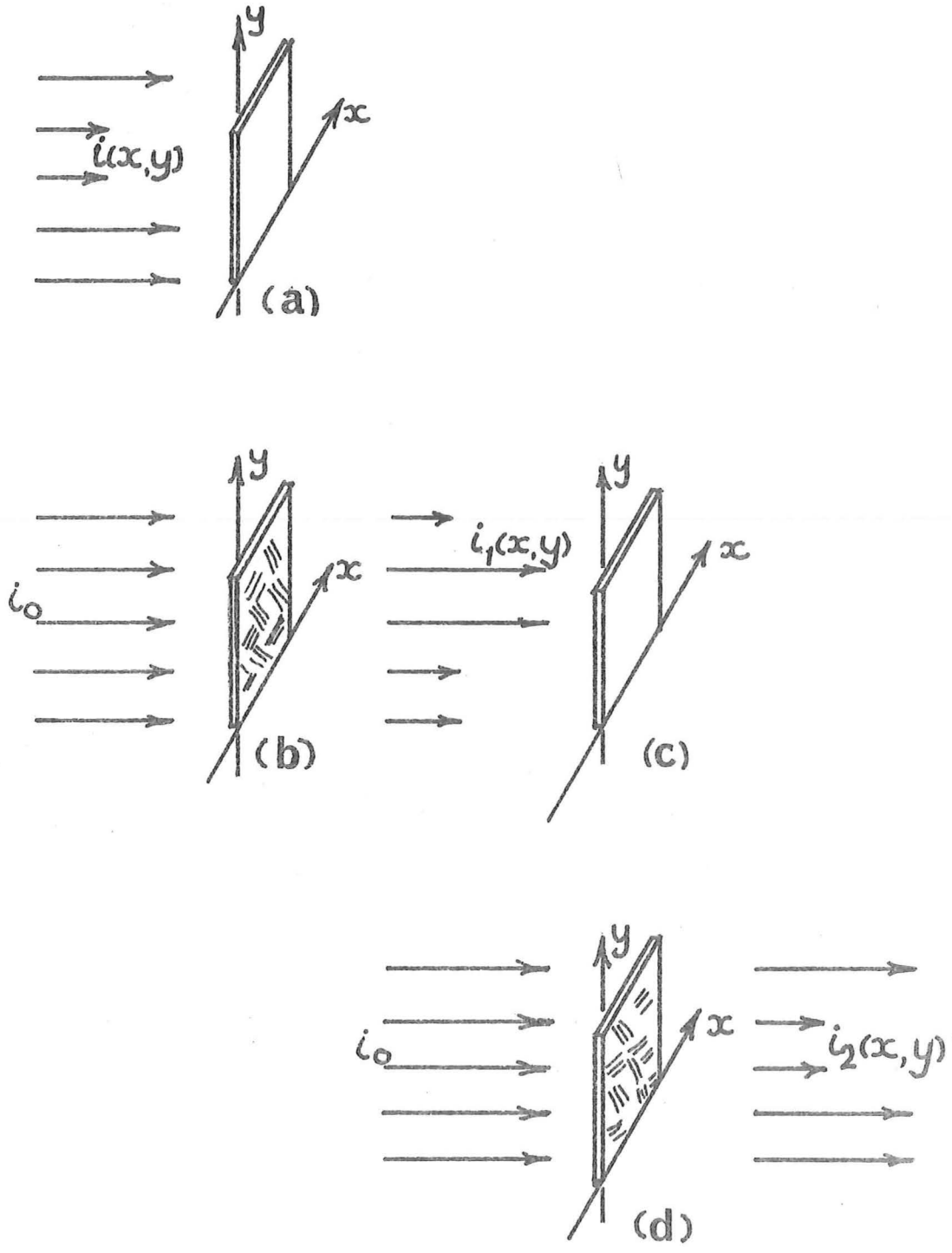


Fig. A5 Producing a positive transparency.
a) Unexposed film b) Developed first negative
c) Unexposed film d) Developed positive.

This ensures that the amplitude distribution of the wavefront transmitted through the positive is proportional to the intensity distribution of the original wavefront. This is necessary since a coherent optical system is linear in amplitude (Goodman, 1968).

A positive with a γ product of 2 was produced in the following way.

A sheet of Panatomic-X was placed in the film plane, and exposed to the image distribution. The exposure time was determined by measuring the average density of the developed negative. If the average density of the developed negative was greater than the median density (Appendix A3) then the film was overexposed, and the exposure time was made shorter to obtain the correct exposure. The film in every case was developed to a γ of 0.95 in Rodinol at a strength of 1/25 for 12 minutes. When a negative with the correct average density, developed to the desired γ was made, then this was reversed onto Copex, developed to a γ of 2.1. The developer was Rodinol at 1/25 for 9 minutes. This positive transparency transmits illumination with an <u>intensity</u> distribution $i_0^2[i(x,y)]^2$ as required, since the overall γ of both steps is approximately 2.0.

APPENDIX B

HOLOGRAPHIC FILM

When a film is used to record a hologram intensity distribution, its characteristics must be considered differently from film used to record an image intensity distribution. Let us first define a new film parameter, transmissivity. This is a measure of the amount of attenuation that a wavefront receives as it passes through a film that has been exposed to uniform illumination, and developed to a specific gamma.

For example, if a wavefront with an amplitude of 1.0 passes through a developed film, and the wavefront immediately behind the film has an amplitude of 0.5, then the transmissivity of that film is 0.5. The transmissivity is related to the density of the film by

$$t = (1/\text{antilog}(d))^2 \qquad \text{(B.1)}$$

Recall that the H-D characteristic curve of film is the range of densities produced by specific levels of exposures. Thus, the H-D curves can be replotted using eqn (B.1) to give the transmissivities produced by a specific level of exposure. This is called the t-E curve of the film. The t-E curve for Kodak 649F used at 6328Å developed in D-19 for five minutes is given in Fig. B1.

Consider a plane sinusoidal diffraction grating of lines that alternate between a transmissivity of greater than 0.5 and less than 0.5. This grating has a constant or bias level

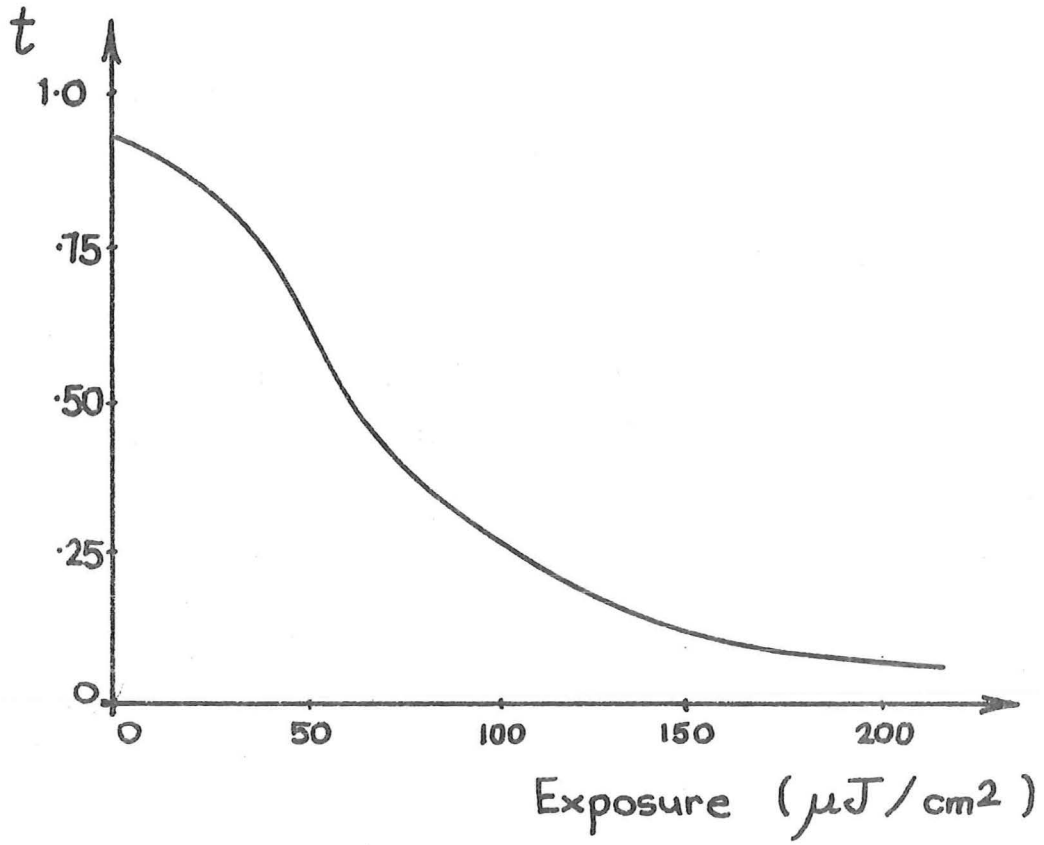


Fig. B1 The t-E curve for Kodak 649F emulsion

of 0.5, and a positive and negative variation around this bias. The grating is illuminated with a quasi-monochromatic beam of light. Behind this grating the light is split into three separate beams. There is a beam that travels straight through the grating and is undeflected. This is usually termed the zero order. There are also beams deflected in equal amounts to the left and right from the zero order.

If the grating is not sinusoidal, or the variation about the bias is not equal, then there may be secondary and even higher orders. The angular deflection of the beams from the zero order is determined by the frequency of the grating and the number of the order.

Now consider a hologram. It consists of a transparency in which there is a grating of fringes variably alternating between a transmissivity of greater than 0.5 and less than 0.5. These fringes are not regular across the whole of the hologram plane since the position and modulation depth of the fringes are effectively a code for the wavefronts emanating from the object. The hologram grating is illuminated with a spatially coherent beam of quasi-monochromatic light. Behind the hologram there is an undeflected beam of light that is the zero order and there are the two beams of light, the primary orders, that form the real and the virtual images.

If the hologram is poorly made, then the grating fringes do not approximate a sinusoid and there may be secondary images due to the secondary orders. The primary and secondary images may overlap so that it is not possible to get a clear image (Goodman, 1968).

Recall that the intensity distribution across a hologram plane is

$$I(\alpha,\beta) = |O(\alpha,\beta) + R(\alpha,\beta)|^2$$

$$= |R|^2 + |O|^2 + RO^* + R^*O \qquad (B.2)$$

Now if $|R| > |O|$ a usual constraint, then

$$I(\alpha,\beta) \doteqdot |R|^2 + RO^* + R^*O \qquad (B.3)$$

where R is a reference wavefront, which usually has a constant amplitude across the α,β plane.

The exposure of the film is the intensity at the film plane multiplied by the exposure time. Thus, the exposure of a film in the hologram plane illuminated for a time ΔT is E which is given by

$$E = I\,\Delta T$$

$$= |R|^2\Delta T + (RO^* + R^*O)\Delta T \qquad (B.4)$$

$R^2\Delta T$ is the constant exposure across the whole hologram plane, and the remaining term in ΔT gives the positive and negative variation around this constant exposure which produces the grating fringes. Consequently for linearity, the portions of the film that receive a constant exposure $|R|^2\Delta T$ when developed should have a transmissivity of 0.5. Referring to the t-E curve of Kodak 649F in the Fig. B1, the exposure time and the amplitude of the reference wavefront are set so that

$$|R|^2\Delta T \doteqdot 60\ \mu\ \text{Joules/cm}^2 \qquad (B.5)$$

The exposure time ΔT that produces a hologram with the correct median transmissivity is usually determined experimentally by exposing a sheet of Kodak 649F to the reference beam alone for an arbitrary exposure time. If the developed film has a transmissivity substantially different from 0.5 (an equivalent density of 0.05) then the exposure time is altered accordingly. When the correct exposure time has been determined, a sheet of the same film is exposed to both the reference and the object wavefronts.

A good rule of thumb way to determine whether a hologram has been formed is to hold the developed film up to a white light source. If a bright multicolour spectrum can be seen through the film, then the film has recorded the hologram distribution.

There are several types of film that may be used to record holographic distributions. Generally we use Kodak 649F spectroscopic plate developed in D-19 (or its equivalent) for five minutes. It resolves up to 5000 lines per mm.

Occasionally we use Agfa Gevaert 10E70 developed in 1/25 Rodinol for five minutes, since this is a faster film and requires only 3 μ Joules/cm^2 to produce a transmissivity level of 0.5. It resolves up to 3000 lines per mm.

Normal photographic film cannot be used for most holographic experiments because of its low resolving power, approximately 200 lines per mm. The resolving power required of the emulsion should be greater than the period between the grating fringes.

APPENDIX C

LIQUID GATES

When a transparency is used as the input to a coherent optical system it is necessary to prevent the variable thickness of the film and backing from distorting the wave-fronts passing through the transparency. This distortion is prevented by placing the transparency between two glass plates whose outside surfaces are polished optically flat to within 1/20 $\bar{\lambda}$. In between the two glass plates and surrounding the transparency is infused a liquid with a refractive index of 1.52. This refractive index represents the best match between the indices of the glass, gelatine and acetate, although in practice some variation in the index of the matching liquid could be tolerated without undue loss of quality.

The main problem with the liquid gate used for experiments reported in this thesis, is the multiple reflections that occur at the rear and front glass-air surfaces. This is illustrated in Fig. C1.

The effect of the multiple reflections is that in the primary focal plane there is always defocused light from the multiple-reflection focal plane.

The multiple reflections may be suppressed by coating the outer faces of the optical flats with a transparent coating of slightly different refractive index to glass and $\bar{\lambda}/4$ thick.

An alternative solution is to suppress the effect of the multiple reflections by making the multiple-reflection focal plane well off to one side. This is done by making one of the glass flats have non-parallel sides. The effect of this can be seen from inspection of the rays in Fig. C2.

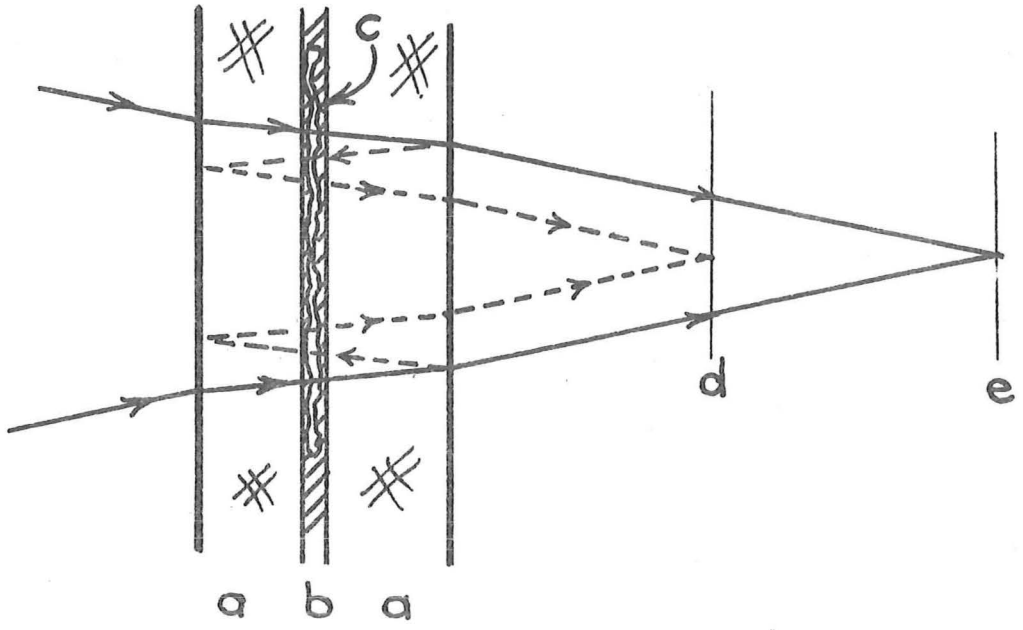


Fig. C1 Liquid gate illustrating multiple reflections.

a) Optical flats b) Index matching liquid
c) Transparency
d) Multiple reflection focal plane
e) Primary focal plane.

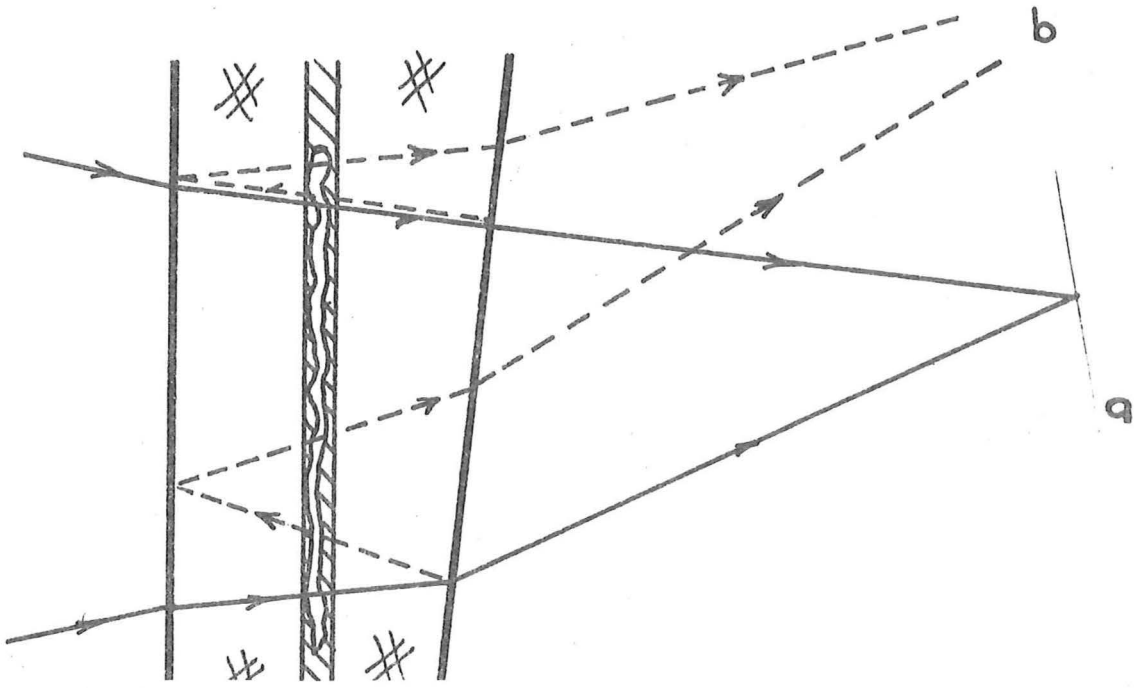


Fig. C2 Non-parallel sided glass flats

a) Primary focal plane
b) Multiple-reflection focal plane.

A disadvantage of this technique is that the optical axis deviates from a straight line. Consequently it is necessary to use a two-dimensional optical bed rather than an optical bench which is restricted to straight line optical systems.

REFERENCES

ABBE, E. (1873) Archiv. Mikroskopische Anat., Vol. 9, 413- .

AGFA (1968) 'Scientia Emulsions for Holography', Gevaert-Agfa N.V., Antwerp.

ANDERSON, G.B. and HUANG, T.S. (1971) "Piecewise Fourier Transformation for Picture Bandwidth Compression", IEEE Trans. Comm. Tech., Vol. COM-19, 133-140.

ANDREWS, H.C. (1970), Computer Techniques in Image Processing Academic Press, New York.

AOKI, Y., BOIVIN, A. (1971) "Hologram Synthesis with Data Derived from a Cross Array", IEEE Trans., AU-19, 312-314.

BATES, R.H.T. and PETERS, T.M. (1971) "Towards Improvements in Tomography", N.Z. J. Sci., Vol. 14, 883-896.

BATES, R.H.T., GOUGH, P.T. and NAPIER, P.J. (1973) "Speckle Interferometry gives Holograms of Multiple Star Systems", Astron. & Astrophys., Vol. 22, 319-320.

BATES, R.H.T., GOUGH, P.T. and PETERS, T.M. (1974) "Some developments in forming Holograms without coherent light", Digest Papers, Intnl. Opt. Compt. Conf., Zurich, Switzerland (IEEE Catalogue 74 CH0862-3C), 100-102.

BATES, R.H.T. and GOUGH, P.T. (1974) "On Optical Processing of Radiation Received from Objects Viewed through Random Media", Submitted to IEEE Trans. Compt. (Special Issue Opt. Computing).

BEARD, T.D. (1969) "Imaging by Correlation of Intensity Fluctuations", Appl. Phys. Let., Vol. 15, 227-229.

BERGER, H. (1969) "A Survey of Ultrasonic Image Detection Methods:, Acoustical Holography, Vol. 1, (Edited by A. Metherell et al.), Plenum Press, New York, 27-48.

BERAN, M.J. and PARRENT, G.B. Jr (1964), Theory of Partial Coherence, Prentice Hall, New Jersey.

BERGER, H. (1969) "A Survey of Ultrasonic Image Detection Methods", Acoustical Holography 1, Plenum Press, New York 27-48.

BERGLAND, G. (1969), "A Guided Tour of the Fast Fourier Transform", IEEE Spectrum, Vol. 6, 41-52.

BOIDIN, D. and LABEYRIE, A. (1974), "Speckle Interferometry in the Photo Counting Mode: m_V = 22 Magnitude Limit Suggested by Theory", Submitted to Ast. Phys. J.

BONNEAU, D. and LABEYRIE, A. (1973) "Speckle Interferometry: Colour Dependent Limb Darkening Evidenced on Alpha Orionis and Omicron Ceti", Ast. Phys. J., Vol. 181, L1-L4.

BORN, M. and WOLF, E. (1959) Principles of Optics, Pergamon Press, London.

BROWN, E.B. (1969) "Prediction and Compensation of Linear Image Motion in Aerial Cameras", NASA, SP-193, 45-51.

BROWN, B.R. and LOHMANN, A.W. (1966) "Complex Spatial Filtering with Binary Masks", App. Opt. Vol. 5, 967-969.

BROWN, B.R. and LOHMANN, A.W. (1969) "Computer Generated Binary Holograms", IBM J. Res. Develop., Vol. 13, 160-168.

BURCKHARDT, C.B. (1970) "A simplification of Lee's method of generating holograms by computer", App. Opt., Vol. 9, 1949.

CHECCACCI, P.F., PAPI, G. and RUSSO, V. (1971) "A Holographic V.H.F. Antenna", IEEE Trans., AP-19, 278-279.

COLLIER, R.J., BURCKHARDT, C.B. and LIN, L.H. (1971), Optical Holography, Academic Press, New York.

CONRADY, A.E. (1929), Applied Optics and Optical Design, Vol. 1, Constable, London.

CONRADY, A.E. (1960) Applied Optics and Optical Design, Vol. 2, Dover, New York.

COOLEY, J.W. and TUKEY, J.W., (1965) "An algorithm for the machine calculation of complex Fourier series", Math. Comput., Vol. 19, 297-301.

CUTRONA, L.J., LEITH, E.N., PORCELLO, L.J. and VIVIAN, W.E. (1966) "On the application of coherent optical processing techniques to synthetic aperture radar", Proc. IEEE, Vol. 54, 1026-1032.

DAINTY, J.C. (1973) "Diffraction-limited imaging of stellar objects using telescopes of low optical quality, Opt. Commun., Vol. 7, 129-134.

DALLAS, W.J. and LOHMANN, A.W. (1972) "Phase quantisation in holograms-depth effects", Ap. Opt. Vol. 11, 192-194.

De ROSIER, D.J. and KLUG, A. (1968) "Reconstruction of three-dimensional structures from electron micrographs", Nature Vol. 217, 130-134.

De VELIS, J.B. and REYNOLDS, G.O. (1967), Theory and Applications of Holography, Addison-Wesley.

DUMONTET, P. (1955) "On Object-Image Correspondence in Optics", Optica Acta, Vol. 2, 53-63.

FRIED, D. (1966) "Optical Resolution through a randomly inhomogeneous medium for very long and very short exposures", J. Opt. Soc. Am., Vol. 56, 1372-1379.

FRASER, D.A.S. (1958) Statistics, An Introduction, John Wiley & Sons, New York.

GABOR, D. (1948) "A new microscope principle", Nature, Vol. 161, 777-778.

GABOR, D. (1949), "Microscopy by reconstructed wavefronts", Proc. Roy. Soc., A197, 454-486.

GABEL, R.A. and LIU, B. (1970) "Minimisation of reconstruction errors with computer generated binary holograms", App. Opt., Vol. 9, 1180-1191.

GEZARI, D.Y., LABEYRIE, A. and STACHNIK, K. (1972), "Speckle Interferometry: Diffraction-Limited measurements of nine stars with the 200-inch telescope", Astrophys. J., Vol. 173, L1-L5.

GOODMAN, J.W. (1968), Introduction to Fourier Optics, McGraw-Hill, New York.

GOODMAN, J.W. (1969) "Digital image formation from detected holographic data", Acoustical Holography, Vol. 1, (Edited by A. Metherell et al.), Plenum Press, New York, 173-185.

GOODMAN, J.W., HUNTLEY, W.H., JACKSON, D.W. and LEHMANN, M. (1966) "Wavefront reconstruction imaging through random media", Appl. Phys. Lett. Vol. 8, 311-313.

GOODMAN, J.W. and SILVESTRI, A.M. (1970) "Some effects of Fourier-domain phase quantization", IBM J. Res. Develop., Vol. 14, 478-484.

GOODMAN, J.W. and KATO, H. (1974) "Logarithmic filtering with coherent optical systems", Digest of Papers, Int. Opt. Comp. Conference, Zurich, Switzerland, IEEE Catalog No. 74 CH-862-3C, pp 31-33.

GOODMAN, J.W. (1969), "Digital Image Formation from Detected Holographic Data", Acoustical Holography 1, Plenum Press, New York, 173-186.

GOUGH, P.T. (1972) "Computation of Accurate Cross-sections of Limbs from Standard X-ray Plates", Proc. 5th Hawaii Int. Conf. on System Sciences, 241-243.

GOUGH, P.T. and BATES, R.H.T. (1972) "Computer generated holograms for processing radiographic data", Comp. Biomed. Res., Vol. 5, 700-708.

GOUGH, P.T. and BATES, R.H.T. (1974), "Speckle Holography", Opt. Act., Vol. 21, 243-254.

GREENE, D. (1969) "Use of acoustical holography for imaging of sources of radiated acoustic energy", J. Acou. Soc. Am., Vol. 46, 44-45.

HARRIS, J.L. (1966), "Image evaluation and restoration", J. Opt. Soc. Am., Vol. 56, 569-574.

HARVEY, J.W. and BRECKINRIDGE, J.B. (1973), "Solar speckle interferometry", Astrophys. J., Vol. 182, L137-L139.

HASKELL, R.E. and CULVER, B.C. (1972) "New Coding Technique for Computer Generated Holograms", App. Opt., Vol. 11, 2712-2714.

HOPKINS, H.H. (1950) Wave Theory of Aberrations, Monographs on the Physics and Chemistry of Materials, Oxford University Press, London.

HOPKINS, H.H. (1966) "The use of diffraction-based criteria of image quality in automatic optical design", Optica Acta, Vol. 13, 343-369.

HORNER, J.L. (1970) "Optical restoration of images blurred by atmospheric turbulence using optimum filter theory", App. Opt., Vol. 9, 167-171.

HOSFELD, R. (1954) "Comparisons of stellar scintillation with image motion", J. Opt. Soc. Am., Vol. 44, 284-288.

HUANG, T.S. (1971) "Digital Holography", Proc. IEEE, Vol. 59, 1335-1346.

HUANG, T.S., SCHREIBER, W.F. and TRETIAK, O.J. (1971) "Image Processing", Proc. IEEE, Vol. 59, 1586-1609.

KENNEDY, W.K. and PETERS, T.M. (1974) "High quality image recording and display using a hybrid computer", Submitted to Aust. Compt. J.

KEETON, S.C. (1968) "A sampled computer-generated binary hologram", Proc. IEEE, Vol. 56 , 325-327.

KING, M.C., NOLL, A.M. and BERRY, D.H. (1970) "A new approach to computer generated holography", Ap. Opt., Vol. 9, 471-475.

KNOWLTON, K. and HARMON, L. (1972) "Computer-produced grey scale", Compt. Graphics & Image Process., Vol. 1, 1-20.

KNOX, K.T. and THOMPSON, B.J. (1973), "New methods of processing speckle pattern star images", Astrophys. J., Vol. 182, L133-136.

KNOX, K.T. and THOMPSON, B.J. (1974) "Recovery of images from atmospherically-degraded short-exposure photographs", Submitted to Ast. Phys. J.

KODAK (1962), Kodak Plates and Films for Science and Industry, Kodak Data Book P-9, Eastman-Kodak Coy., Rochester, N.Y.

KOHLER, D. and MANDEL, L. (1973), "Source reconstruction from the modulus of the correlation function: a practical approach to the phase problem in optical coherence theory", J. Opt. Soc. Am., Vol. 63, 126-134.

KOLOGELNIK, H. (1965) "Holographic image projection through inhomogeneous media", Bell System Tech. J., Vol. 44, 2451-2455.

KORFF, D. (1973) "Analysis of a method for obtaining near-diffraction-limited information in the presence of atmospheric turbulence", J. Opt. Soc. Am.,, Vol. 63, 971-980.

KORFF, D., DRYDEN, G. and MILLER, M.G. (1972) "Information retrieval from atmospheric induced speckle patterns", Opt. Commun., Vol. 5, 187-192.

LABEYRIE, A. (1970), "Attainment of diffraction limited resolution in large telescopes by Fourier analysing speckle patterns in star images", Astron. and Astrophys. Vol. 6, 85-89.

LABEYRIE, A (1974) Private communication.

LEE, W.H. (1970) "Sampled Fourier transform hologram generated by computer", Ap. Opt., Vol. 9, 639-643.

LEITH, E.N. and UPATNIEKS, J. (1963) "Wavefront reconstruction with continuous-tone objects", J. Opt. Soc. Am., Vol. 53, 1377-1381.

LEITH, E.N. and UPATNIEKS, J. (1964) "Wavefront reconstruction with diffused illumination and three-dimensional objects", J. Opt. Soc. Am., Vol. 54, 1295-1301.

LEITH, E.N. and UPATNIEKS, J. (1966) "Holographic imagery through diffusing media", J. Opt. Soc. Am., Vol. 56, 523.

LESEM, L.B., HIRSCH, P.M. and JORDAN, J.A. Jr (1968), "Computer synthesis of holograms for 3-D display", Com. ACM, Vol. 11, 661-674.

LIU, C.Y.C. and LOHMANN, A.W. (1973) "High resolution image formation through the turbulent atmosphere", Opt. Comm, Vol. 8, 372-377.

LOHMANN, A.W. and PARIS, D.P. (1967) "Binary Fraunhofer holograms, generated by computer", App. Opt., Vol. 6, 1739-1748.

LOHMANN, A.W. and PARIS, D.P. (1968) "Computer generated spatial filters for coherent optical data processing", App. Opt. Vol. 7, 651-655.

LOWENTHAL, S. and CHAVEL, P. (1974) "Reduction of the number of samples in computer holograms for image proce-sing", App. Opts., Vol. 13, 719-720.

MacADAM, D.P. (1970) "Digital image restoration by constrained decomvolution", J. Opt. Soc. Am., Vol. 60, 1617-1627.

McGLAMERY, B.L. (1967) "Restoration of turbulence-degraded images", J. Opt. Soc. Am., Vol. 57, 293-297.

MANDEL, L. and WOLF, E. (1961) "Some properties of coherent light", J. Opt. Soc. Am., Vol. 51, 815-819.

MARÉCHAL, A. and CROCE, P. (1953) "Un filtre de fréquences spatiales pour l'amélioration due contraste des images optiques", Compt. Rendus, Vol. 237, 607.

MEHTA, C.L. (1968) "New approach to the phase problem in optical coherence theory", J. Opt. Soc. Am., Vol. 58, 1233-1234.

MEYER, A.J. and HICKLING, R. (1967) "Holograms synthesised on a computer-operated cathode-ray tube", J. Op. Soc. Am., Vol. 57, 1388-1389.

MILLER, M.G., SCHNEIDERMAN, A.M. and KELLEN, P.F. (1973) "Comparison of methods for processing short-exposure data from large telescopes", Astrophys. J., Vol. 186, L91-L93.

MUELLER, P.F. and REYNOLDS, G.O. (1967), "Image restoration by removal of random-media degradations", J. Opt. Soc. Am. Vol. 57, 1338-1344.

NAPIER, P.J. and BATES, R.H.T. (1971) "Holographic approach to radiation pattern measurement - II. Experimental verification", Int. J. Engng. Sci., Vol. 9, 1193-1208.

NAPIER, P.J. and BATES, R.H.T. (1973) "Antenna aperture distributions from holographic type of radiation - Pattern measurement", Proc. IEE, Vol. 120, 30-34.

NASA (1969) "Evaluation of motion degraded images", NASA, SP-193.

OPPENHEIM, A.V., SCHAFER, R.W. and STOCKHAM, T.G. Jr (1968) "Nonlinear filtering of multiplied and convolved signals", Proc. IEEE, Vol. 56, 1264-1291.

PAPOULIS, A. (1968) Systems and Transforms with Applications in Optics, McGraw-Hill, New York.

PERRY, B. and MENDELSON, M.L. (1964) "Pictorial output with a standard line printer", Comm. A.C.M., Vol. 7, 311-313.

PRATT, W.K., KANE, J. and ANDREWS, H.C. (1969) "Hadamard transform in image coding", Proc. IEEE, Vol. 57, 439-441.

ROBBINS, G.M. and HUANG, T.S. (1972) "Inverse filtering for linear-shift- variant imaging systems", Proc. IEEE, Vol.60, 862-871.

RUSSEL, F.D. and GOODMAN, J.W. (1971) "Nonredundant arrays and post-detection processing for aberration compensation in incoherent imaging", J. Opt. Soc. Am., Vol. 61, 182-191.

SMITH, P.R., PETERS, T.M. and BATES, R.H.T. (1973) "Image reconstruction from finite numbers of projection", J. Phys. A. Vol. 6, 361-382.

STROKE, G.W. (1969) "Image deblurring and aperture synthesis using posteriori processing by Fourier-transform holography", Opt. Act., Vol. 16, 401-422.

STROKE, G.W. and HALIOUA, M. (1972) "Attainment of diffraction-limited imaging in high-resolution electron microscopy by 'a Posteriori' holographic image sharpening, I", Optik, Vol. 35, 50-65.

TAUB, H. and SCHILLING, D.L. (1971) Principles of Communications Systems, McGraw-Hill (New York).

TEXEREAU, J. (1963), "Limitations à la qualité des iamges d'un grand telescope", App. Opt., Vol. 2, 23-30.

TODD, H.N. and ZAKIA, R.D. (1969), Photographic Sensitometry. The Study of Tone Reproduction, Morgan and Morgan, New York.

TRICOLES, G. and ROPE, E.L. (1967) "Reconstruction of visible images from reduced scale replicas of microwave holograms", J. Opt. Soc. Am., vol. 57, 97-99.

UPATNIEKS, J., VANDER LUGT, A. and LEITH, E.N. (1966), Correction of lens aberrations by means of holograms, Appl. Opt., Vol. 5, 589-593.

VANDER LUGT (1964) "Signal detection by complex spatial filtering", IEEE Trans. Inform. Theory, Vol. IT-10, 139-145.

WANG, C.P. (1974) "High resolution imaging through turbulent media", Opt. Comm., Vol. 10, 253-257.

WATERS, J.P. (1966) "Holographic image synthesis utilising theoretical methods", App. Phys. Let., Vol. 9, 405-407.

WATERS, J.P. (1968) "Three-dimensional Fourier-transform method for synthesizing binary holograms", J. Op. Soc. Am., Vol. 58, 1284-1288.

WATERS, J.P. and MICHAEL, F. (1969) "High resolution images from C.R.T.-generated synthetic holograms", App. Opt., Vol. 8, 714-715.

WYNNE, G.C. (1959) "Lens designing by electronic computer - I", Proc. Phys. Soc. of Lond., Vol. 73, 777-787.

ZERNIKE, F. (1935) "Das Phasenkontrastverfahren bei der Mikroskopischen Beobachtung", Z. Tech. Phys., Vol. 16, 454-463.